分类号 O572.21 密级 公开

UDC 530

学 位 论 文

ATLAS 实验中通过 $b\bar{b}\tau^{+}\tau^{-}$ 衰变道寻找希格斯玻色子的对产生

（题名和副题名）

张博文

（作者姓名）

指导教师姓名、职务、职称、学位、单位名称及地址

陈申见 教授 博士 南京大学 汉口路 22 号

申请学位级别 博士 专业名称 粒子物理与原子核物理

论文提交日期 2022 年 5 月 26 日 论文答辩日期 2022 年 5 月 19 日

学位授予单位和日期

答辩委员会主席： 刘衍文 教授

评阅人： 孙小虎 教授

李数 教授

肖朦 教授

张雷 教授

2022 年 5 月 26 日

南京大学

NANJING UNIVERSITY

研 究 生 毕 业 论 文

（申请博士学位）

论 文 题 目	ATLAS 实验中通过 $b\bar{b}\tau^{+}\tau^{-}$ 衰变道寻找希格斯玻色子的对产生
作 者 姓 名	张博文
专 业 名 称	粒子物理与原子核物理
研 究 方 向	粒子物理实验
指 导 教 师	陈申见 教授

2022 年 5 月 26 日

学　　号：DZ1722040

论文答辩日期：2022 年 5 月 19 日

指 导 教 师：　　（签字）

Search for Higgs boson pair production in the $b\bar{b}\tau^{+}\tau^{-}$ decay channel with the ATLAS experiment

by

Zhang Bowen[1]

Directed by

Professor Chen Shenjian[1]

[1]School of Physics, Nanjing University

May,2022

Submitted in partial fulfilment of the requirements
for the degress of PhD in Particle Physics and Nuclear Physics

南京大学研究生毕业论文中文摘要首页用纸

毕业论文题目：ATLAS 实验中通过 $b\bar{b}\tau^+\tau^-$ 衰变道寻找希格斯玻色子的对产生

粒子物理与原子核物理 专业 2017 级博士生姓名： 张博文

指导教师（姓名、职称）： 陈申见 教授

标准模型中，希格斯玻色子的自耦合通过电弱对称性破缺机制引入，其作用强度与希格斯玻色子的质量以及弱相互作用耦合常数相关。在大型强子对撞机（LHC）的实验环境下，希格斯玻色子对（HH）主要通过希格斯玻色子的自耦合和其与顶夸克的汤川耦合产生，这两种过程的干涉压低了希格斯玻色子的对产生截面，使其约为单希格斯玻色子产生截面的千分之一。许多新物理模型中，新标量粒子或反常耦合的引入可以改变希格斯势的形式和希格斯玻色子自耦合的强度，进而提高 HH 的产生率，使得实验上有可能以目前采集的数据观测并发现希格斯玻色子的对产生过程。寻找希格斯玻色子的对产生是自 2012 年希格斯玻色子发现之后近年来 ATLAS 实验的重点研究课题之一，这方面的研究可以增进我们对希格斯玻色子自耦合和希格斯势性质的认识，并加深对标准模型电弱理论以及早期宇宙热力学演化的理解。

本文利用 ATLAS 探测器采集的 LHC 上质心系能量为 13 TeV，积分亮度为 139 fb^{-1} 的质子–质子对撞数据，通过含有两个底夸克喷注和两个陶轻子的事例寻找希格斯玻色子的对产生过程，主要研究两个陶轻子全部衰变到强子的衰变道（$\tau_{\text{had}}\tau_{\text{had}}$），同时给出此衰变道和半轻衰变道（$\tau_{\text{lep}}\tau_{\text{had}}$）统计学联合分析的结果。根据末态粒子的拓扑结构，分析分为两类：解析信号分析（称作 Resolved 分析）和融合信号分析（称作 Boosted 分析）。

在解析信号分析中，两个底夸克喷注和两个陶轻子可以被分别重建。此分析考虑了非共振态和共振态这两种 HH 产生模式。其中非共振态产生模式的信号特征基于标准模型胶子–胶子融合（ggF）和矢量玻色子融合（VBF）产生模式的信号特征，而共振态产生模式下，HH 通过重的、窄宽度的标量粒子衰变产生，考虑的共振态粒子质量范围在 1.6 TeV 以内。在对非共振态产生模式的寻找中未观测到显著的超出标准模型本底预期的迹象。只考虑 $\tau_{\text{had}}\tau_{\text{had}}$ 衰变道的情况下，观测（预期）非共振态 HH 产生截面在 95% 置信度下的上限是标准模型预期 ggF+VBF 产生截面的 5.0（4.4）倍，联合 $\tau_{\text{lep}}\tau_{\text{had}}$ 衰变道后相应结果为 4.7（3.9）倍。另外，非共振态产生模式的分析还对希格斯玻色子的自耦合强度进行了约束，在 95% 的置信度下将自耦合强度修饰因子 κ_λ 的观测（预期）取值范围限制在 $[-2.5, 9.7]$（$[-2.3, 9.8]$）之内。在寻找共振态产生模式的分析中，观测到的最显著偏差位于共振态粒子质量为 1 TeV 处，对应的局部（整体）显著性在只考虑 $\tau_{\text{had}}\tau_{\text{had}}$ 衰变道的情况下为 2.8σ（1.7σ），相应的联合分析结果为 3.0σ（2.0σ）。在 95% 置信度下，只考虑 $\tau_{\text{had}}\tau_{\text{had}}$ 衰变道的观测（预期）共振态 HH 产生截面上限为 27–1600 fb（18–1000 fb），联合分析的结果为 23–930 fb（12–840 fb），具体取值需参考共振态粒子的质量假设。

在融合信号分析中，两个较窄的底夸克喷注融合成单个宽喷注，两个强子衰变的陶轻子融合成单个双陶体（称作 di-τ），只有融合后的物理对象可以被重建。此分析着重于

寻找通过共振态模式产生的 HH，应用双陶体标记方法将共振态粒子质量研究范围扩展到了 3 TeV。最终，未寻找到显著的超出标准模型本底预期的迹象，给出观测（预期）共振态 HH 产生截面在 95% 置信度下的上限，结果为 28–94 fb（32–74 fb），具体取值需参考共振态粒子的质量假设。

此外，本文还提出了一种崭新的陶轻子强子型衰变模式分类方法。这种分类方法基于 DeepSet 神经网络架构，其分类效率和分类纯度均比当前方法提高了百分之九，此新方法有望应用在未来的物理分析中。

关键词： ATLAS 实验；希格斯玻色子对产生；希格斯玻色子自耦合；陶轻子；底夸克

南京大学研究生毕业论文英文摘要首页用纸

THESIS: Search for Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel with the ATLAS experiment

SPECIALIZATION: Particle Physics and Nuclear Physics

POSTGRADUATE: Zhang Bowen

MENTOR: Professor Chen Shenjian

In the Standard Model, the self-coupling of the Higgs boson is introduced by the electroweak symmetry breaking mechanism, the strength of which is related to the mass of Higgs boson and the coupling constant of weak interaction. At the Large Hadron Collider (LHC), Higgs boson pairs (HH) are dominantly produced through the self-coupling of Higgs bosons and the Yukawa coupling with top-quarks. The interference of these two processes suppresses the production cross section of the HH to about one thousandth of that of the single Higgs boson. Nevertheless, many new physics models predict new scalar particles or anomalous couplings that can change the form of the Higgs potential and the strength of the self-coupling, thereby increasing the yield of HH events, and making it possible to observe Higgs boson pair production with the data collected up to the present. The search for Higgs boson pair production is one of the key research topics of the ATLAS experiment in recent years since the discovery of the Higgs boson in 2012. It can improve our knowledge of the nature of the Higgs boson self-coupling and the Higgs potential, as well as deepen our understanding of the electroweak theory of the Standard Model and the thermal history of the early universe.

In this thesis, a search for the Higgs boson pair production is performed in events with two b-jets and two τ-leptons, by exploring the proton-proton collision data collected by the ATLAS detector at the LHC with a centre-of-mass energy of 13 TeV and an integrated luminosity of 139 fb^{-1}. The $\tau_{\text{had}}\tau_{\text{had}}$ decay channel in which all τ-leptons decay to hadrons is studied in detail, and the results are statistically combined with a semi-leptonic decay channel called the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. Depending on the topology of the final-state physics objects, the analysis is divided into two types — the resolved and the boosted (a.k.a merged) analyses.

In the resolved analysis, the two b-jets and two τ-leptons can be reconstructed individually. Both the non-resonant and resonant HH production modes are considered in this analysis. The signatures of the non-resonant mode are based on those of the Standard Model gluon-gluon fusion (ggF) and vector boson fusion (VBF) modes, while in the resonant production mode, the Higgs boson pairs are produced by the decay of heavy, narrow width, scalar particles, with mass below 1.6 TeV. In the search for non-resonant production mode, no significant excesses of events are observed above the expectations of the Standard Model background. Observed (expected) upper limits are placed at 95% confidence-level on the non-resonant HH production cross section,

which are 5.0 (4.4) times the expected ggF+VBF HH production cross section considering only the $\tau_{\text{had}}\tau_{\text{had}}$ channel, and the corresponding results are 4.7 (3.9) after combining the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. In addition, observed (expected) constraints are set on the strength of Higgs boson self-coupling in the analysis of non-resonant production mode. The values of self-coupling modifier κ_λ that lie outside $[-2.5, 9.7]$ ($[-2.3, 9.8]$) are excluded at 95% confidence-level. The most significant deviation is observed in the search for resonant production mode at a resonance particle mass of 1 TeV, corresponding to a local (global) significance of 2.8σ (1.7σ) considering only the $\tau_{\text{had}}\tau_{\text{had}}$ channel, and 3.0σ (2.0σ) after combining the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. At 95% confidence-level, the observed (expected) upper limits on the cross section of resonant HH production are 27–1600 fb (18–1000 fb) considering only the $\tau_{\text{had}}\tau_{\text{had}}$ channel, and 23–930 fb (12–840 fb) after the combination. The actual value depends on the mass hypothesis of the resonance particle.

In the boosted analysis, the two small radius jets merge into a single large radius jet, and the hadronic decay products of two τ's merge into a single object called di-τ, thus only the merged objects are reconstructed. This analysis focuses on the resonant HH production, and it extends the search range of resonance mass to 3 TeV by using the di-τ tagging technique. Finally, no significant excesses of events are observed with respect to the Standard Model background prediction. The observed (expected) upper limits on the cross section of resonant HH production are placed, which are 28–94 fb (32–74 fb), depending on the mass hypothesis of the resonance particle.

Besides, a novel method is proposed to classify the hadronic decay modes of τ-leptons. The classification method based on a DeepSet neural network architecture improves both the classification efficiency and purity by nine percent compared with current method. It is promising to apply this new method in future physics analyses.

KEYWORDS: ATLAS experiment; Higgs boson pair production; Higgs boson self coupling; tau lepton; bottom quark

Contents

List of Figures

List of Tables

Introduction

After the discovery of the Higgs boson [1, 2], the properties of Higgs boson have been extensively studied by the ATLAS and CMS experiment at the Large Hadron Collider (LHC). Great achievements have been made on the measurement of couplings between Higgs boson and other Standard Model (SM) particles [3, 4]. In particular, the Yukawa couplings between Higgs boson and τ-leptons as well as bottom-quarks have been observed [1]. However, we still lack knowledge about the self-coupling of Higgs boson. The self-coupling of the Higgs boson originates from the Higgs potential by the electroweak symmetry breaking mechanism. In the SM, the self-coupling strength is constrained by the mass of the Higgs boson and the coupling constant of weak interaction. The Higgs potential is the central element in the electroweak theory of the SM, and also a key ingredient in the thermal history of the early universe, which is related to fundamental questions such as the stability of electroweak vacuum and the nature of electroweak phase transition. The study of self-coupling can improve our understanding of these questions.

The LHC uniquely provides an opportunity to determine the self-coupling of the Higgs boson, that is, to measure the cross section of Higgs boson pair production. At the LHC, Higgs boson pair (HH) can be produced by the decay of a virtual Higgs boson in trilinear Higgs boson self-coupling. This production mode is interfered by the production mode that involves two top-quark Yukawa couplings. The total production cross section is suppressed by the interference to about 10^{-3} times that of the single Higgs boson [5], thus the HH process has not been observed so far. Nevertheless, the magnitude of Higgs boson pair production cross section is very sensitive to the presence of physics beyond the SM (BSM). In various BSM theories, the production rate of Higgs boson pair can be enhanced either by anomalous couplings or by resonant effects [6]. The search for non-resonant and resonant HH production modes aims to find these signals, respectively.

Several HH decay channels have been explored in the search for Higgs boson pair production in early Run-2 data-taking period of the LHC [7, 8], using the proton-proton collision data of an integrated luminosity of 36.1 fb^{-1}. It is found that the most sensitive decay channels are $b\bar{b}b\bar{b}$, $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^+\tau^-$. This thesis focuses on the $b\bar{b}\tau^+\tau^-$ decay channel, in which the signal events contain two b-jets and two τ-leptons, each from the decay of the Higgs boson. This decay channel benefits from relatively less background contamination than the $b\bar{b}b\bar{b}$ channel, and higher branching ratio

[1] The observed $Hb\bar{b}$ and $H\tau\tau$ couplings are compatible with the SM prediction. The conclusion is important for this thesis, in which the branching ratios of $H \to bb$ and $H \to \tau\tau$ are assumed to be their SM values.

than the $b\bar{b}\gamma\gamma$ channel. Consequently, the $b\bar{b}\tau^+\tau^-$ channel is a very competitive decay channel in the search for HH production. However, the $b\bar{b}\tau^+\tau^-$ channel is also very challenging to be analysed. Because of the energy carried away by the neutrinos in τ-lepton decays, the signature of signal event is not as clear as the other channels. Moreover, due to the decay of τ-leptons and the hadronic final state, multitudinous background processes need to be considered. In the search for non-resonant HH production, an enlargement of more than 13 times the SM HH production cross section has been excluded at 95% confidence-level by the ATLAS experiment [9], and the corresponding result was 30 by the CMS experiment [10]. In the search for resonant HH production, 95% confidence-level upper limit on the HH production cross section via a heavy resonance particle has been determined. Although different benchmark signals were used in ATLAS and CMS experiments, the mass search range did not exceed 1 TeV.

The Run-2 data-taking has been completed in 2018 successfully, during which, the ATLAS detector collected an integrated luminosity of 139 fb^{-1} proton-proton collision data for physics studies. Meanwhile, significant improvements have been made in the reconstruction and identification performance of b-jets and hadronic τ-leptons, and new techniques have been developed to reconstruct boosted (a.k.a merged) objects. The above advantages not only bring a sizeable gain in experimental sensitivity, but also open up the opportunity to explore higher energy regions.

The analysis contains $\tau_{\text{had}}\tau_{\text{had}}$ and the $\tau_{\text{lep}}\tau_{\text{had}}$ decay channels based on the hadronic or leptonic decay products of the τ-lepton pair. The $\tau_{\text{had}}\tau_{\text{had}}$ channel is studied in-depth in this thesis. According to the topology of the final-state physics objects, the analysis is divided into two types — the resolved and the boosted analyses.

In the resolved analysis, the two b-jets and decay products of two τ-leptons can be reconstructed individually. The dominant backgrounds are from vector boson in association with jets, single Higgs boson, multi-jet, and top-quark processes. Multivariate analysis (MVA) based on the invariant masses and angular relations of the final state objects is performed to extract the HH signal events. The outputs of the MVA methods are used as the final signal-background discriminant variable. The results from the $\tau_{\text{had}}\tau_{\text{had}}$ channel are statistically combined with those from the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. Both the non-resonant and resonant HH production modes are searched. Finally, the sensitivity is improved by a factor of about four in both HH production modes compared with the sensitivity of that in early Run-2.

In the boosted analysis, the two small radius jets merge into a single large radius jet, and the hadronic decay products of two τ's merge into a single object called di-τ [2], thus only the merged objects are reconstructed. It is the first time in ATLAS physics analysis that the hadronic boosted di-τ tagging technique is utilised. The vector boson in association with jets, single Higgs boson, and multi-jet processes are the major background processes. A cut-based strategy is used to select the signal candidates. The results are extracted by counting the number of remaining events after applying the selections. The boosted analysis only searches for the resonant HH production, and

[2] Only the fully-hadronic di-τ is considered.

it extends the resonance mass search range to 3 TeV.

Structure of the thesis

The thesis is composed as follows. The theoretical background of the SM and Higgs boson pair production is introduced in Chapter 1, followed by an overview of the LHC and the ATLAS detector in Chapter 2. Definitions of the physics objects that are used in the physics analysis (excepting the hadronic τ-lepton) and a general description of performing physics analysis in ATLAS experiment are also described in the same chapter. Chapter 3 introduces the reconstruction and identification of hadronically decaying τ-leptons, with focus on the new decay mode classification method. The searches for Higgs boson pair production in resolved and boosted $b\bar{b}\tau^+\tau^-$ final states are presented comprehensively in Chapter 4 and 5. Lastly, Chapter 6 encapsulates the thesis and prospects the future studies.

Clarification of personal contributions

The ATLAS experiment is run by a collaboration of about 3000 scientists and students. Due to the complexity of the experiment, the work is usually shared by a group of people. Hence, it is useful to clarify my personal contributions.

Boosted analysis of HH search

I worked on most parts of the analysis, including the background estimations of multi-jet, Z boson in association with heavy flavour jets, optimisation of sensitivity, event categorisation, and statistical analysis. I also worked on the development of analysis software and the production of slimmed dataset. I was one of the editors of the internal documentation. I edited the HEPData entry of this analysis. I was the statistics liaison to the Di-Higgs physics group, for the combination of all HH decay channels, and I presented this analysis to the ATLAS collaboration on behalf of the analysis team for the approval of paper publication.

Resolved analysis of HH search

People in this analysis team worked on either the $\tau_{\text{lep}}\tau_{\text{had}}$ channel or the $\tau_{\text{had}}\tau_{\text{had}}$ channel, and a few of them worked on the combination of these two channels. I worked on the (full-top-mass) generator validation of ggF HH signal, the estimation of fake-τ_{had} $t\bar{t}$ and Z boson with association of heavy flavour jets background processes, and the theoretical uncertainty of major backgrounds, also worked on the development of analysis software and the production of slimmed dataset. I developed the $t\bar{t}$ background re-weighting method, which is also applied in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. I

was in charge of the joint statistical analysis of the $\tau_{\text{had}}\tau_{\text{had}}$ and the $\tau_{\text{lep}}\tau_{\text{had}}$ channels, and I presented the whole analysis to the ATLAS collaboration for the approval of paper publication.

Other contributions to the ATLAS experiment

I studied the performance of τ_{had} reconstruction and identification at high-luminosity LHC environment. I worked on the software for the tuning and testing of τ_{had} reconstruction and identification algorithms in the tau performance working group. I developed the DeepSet τ_{had} decay mode classification method. I served as the machine learning liaison between the tau performance working group and ATLAS machine learning forum. I worked on monitoring the operation of the ATLAS distributed computing system.

Chapter 1

Theoretical Background

This chapter gives an overview of the theoretical framework of particle physics with a focus on the physics of the Higgs boson in the context of the LHC. Furthermore, the motivation to search for Higgs boson pair production is explained. The main reference literature of this chapter includes Ref. [11–13].

1.1 The Standard Model

1.1.1 Fundamental particles and interactions

The Standard Model (SM) [14–16] of particle physics is a theoretical framework that provides a fundamental description of the electromagnetic, weak and strong interactions in a unified picture, together with the elementary particles. In the past decades, it is extensively tested by experiments. The results are dominantly in line with the prediction of the SM. It makes SM the most remarkable triumph of particle physics. The SM is constructed based on Quantum Field Theory (QFT) that describes the elementary particles and their interactions, the local gauge principle that determines the nature of these interactions, and the mechanism of Electroweak (EW) symmetry breaking that gives masses to the particles.

The fundamental interactions are constructed in a $SU(3)_C \times SU(2)_L \times U(1)_Y$ symmetry group, which gives rise to twelve vector fields. The excitations of these vector fields are known as gauge bosons that propagate the interactions (forces). These gauge bosons are

- Eight bi-coloured gluons (g) that propagate the strong interaction;
- $W^{\pm}$ and Z bosons that propagate the weak interaction;
- Photon (γ) that propagates the electromagnetic interaction.

The dynamics of the spinor fields is ruled by the Dirac equation. The excitations of the spinor fields are the building blocks of the matter world, called matter particles or fermions. These fermions are categorised into leptons and quarks by their interaction behaviours, which are

- Leptons that do not participate in strong interaction, labelled as $\binom{\nu_e}{e}, \binom{\nu_\mu}{\mu}, \binom{\nu_\tau}{\tau}$;

- Quarks that participate in both EW and strong interactions, labelled as $\binom{u}{d}, \binom{c}{s}, \binom{t}{b}$.

As shown above, there are six leptons and six quarks in the SM, and they are organised into three generations by their masses and decays. Ultimately, in the electroweak symmetry breaking mechanism, scalar fields are required to generate mass to the bosons and fermions. The Higgs boson (H) arises from this mechanism.

The interactions can be split into two parts — the strong interaction and the joint EW interaction. They are briefly reviewed below. Following the formalism of the QFT, the dynamic and interaction terms of the quantum fields are described by Lagrangian densities [1].

1.1.1.1 Strong interaction

The strong interaction between quarks and gluons is described by the Quantum Chromodynamics (QCD). The QCD is a gauge theory that is invariant under $SU(3)$ local gauge transformation in the colour space. The symmetry group is called $SU(3)_C$. The Lagrangian of QCD can be schematically written as

$$\mathscr{L}_{\text{QCD}} = \bar{Q} i\gamma^{\mu} D_{\mu} Q - \frac{1}{4} G^{a}_{\mu\nu} G^{\mu\nu}_{a}, \tag{1.1}$$

where $Q(x)$ is the quark field. The notation $Q(x)$ represents that the quark field Q is a function of the four-dimensional space-time x. Similar notation is used throughout this chapter. Eight massless vector fields, the gluon fields G^{a}_{μ} are introduced to satisfy the requirement of gauge invariance. They form a covariant derivation

$$D_{\mu} = \partial_{\mu} + i g_S T_a G^{a}_{\mu} \tag{1.2}$$

and a gluon tensor field

$$G^{a}_{\mu\nu} = \partial_{\mu} G^{a}_{\nu} - \partial_{\nu} G^{a}_{\mu} - g_S f_{abc} G^{b}_{\mu} G^{c}_{\nu}, \tag{1.3}$$

where f_{abc} and $T_a (a = 1 \ldots 8)$ are the structure constants and generators of $SU(3)_C$. The generators T_a can be conventionally represented by the Gell-Mann matrices $\frac{\lambda_a}{2}$. The strong interaction constant g_S, more widely known as the strong interaction strength $\alpha_S = \frac{g_S^2}{4\pi}$, depends on the energy scale of the interaction. At high energy, α_S is enough small that perturbation calculations become valid. This property is known as asymptotic freedom.

Free quarks are not experimentally observable due to the hypothesis of colour confinement that originates from the gluon self-interactions in QCD. Quarks must be confined into colour singlet states (a.k.a colourless states) called hadrons. High energy quarks and gluons can produce jets of hadrons by a process including parton shower and hadronisation. Phenomenological models are used to simulate the picture of this process from the scale of hard process to the scale of hadronisation. They are usually used within parton shower event generators.

[1] For clarity, Lagrangian is Lagrangian density in the following.

1.1.1.2 Electroweak interaction

Electromagnetic interaction and weak interaction are unified in a joint gauge theory associated with $SU(2)_L \times U(1)_Y$ symmetry group. The first part comes from the $SU(2)$ symmetry of the weak charged-current interaction, where L indicates that the weak charged-current interaction only couples to left-handed chiral (LH) particle states and right-handed chiral (RH) antiparticle states. The unification with the second part incorporates the weak neutral-current interaction with the electromagnetic interaction, where Y stands for weak hypercharge. The schematic Lagrangian of the EW interaction can be written as

$$\mathscr{L}_{\mathrm{EW}} = \bar{L} i\gamma^{\mu} D_{\mu} L + \bar{R} i\gamma^{\mu} D_{\mu} R - \frac{1}{4} W^{a}_{\mu\nu} W^{\mu\nu}_{a} - \frac{1}{4} B_{\mu\nu} B^{\mu\nu}, \tag{1.4}$$

where $L(x) = \frac{1}{2}(1-\gamma^5)F(x)$ denotes the LH spinor (fermion) field that can be termed as weak isospin doublet

$$\begin{pmatrix} \nu_e \\ e \end{pmatrix}_L, \begin{pmatrix} \nu_\mu \\ \mu \end{pmatrix}_L, \begin{pmatrix} \nu_\tau \\ \tau \end{pmatrix}_L, \begin{pmatrix} u \\ d' \end{pmatrix}_L, \begin{pmatrix} c \\ s' \end{pmatrix}_L, \begin{pmatrix} t \\ b' \end{pmatrix}_L. \tag{1.5}$$

The term $F(x)$ represents the generic fermion field. The prime on the superscripts of the quark doublet indicates that the spinor field is in weak interaction eigenstates. The transformation from mass eigenstate to weak interaction eigenstate is accomplished by the Cabibbo–Kobayashi–Maskawa (CKM) matrix for quarks and Pontecorvo-Maki-Nakagawa-Sakata (PMNS) matrix for neutrinos. $R(x) = \frac{1}{2}(1+\gamma^5)F(x)$ denotes the RH spinor field which is a weak isospin singlet

$$e^-_R, \mu^-_R, \tau^-_R, u_R, c_R, t_R, d_R, s_R, b_R. \tag{1.6}$$

There is no RH neutrino or LH anti-neutrino in the SM. Note that the covariant derivative in Equation (1.4) acts differently on L and R due to the V-A structure of the weak interaction. They are

$$\begin{aligned} D_{\mu} L &= [\partial_{\mu} + i g' \frac{Y}{2} B_{\mu} + i g T_a W^a_{\mu}] L \\ D_{\mu} R &= [\partial_{\mu} + i g' \frac{Y}{2} B_{\mu}] R. \end{aligned} \tag{1.7}$$

In the following equations, g' and g are associated with the strength of the $U(1)$ and $SU(2)$ couplings, respectively [2]. The vector fields $W^a_{\mu}(a = 1, 2, 3)$ and B_{μ} are introduced to ensure the gauge symmetry. They form tensor fields

$$\begin{aligned} B_{\mu\nu} &= \partial_{\mu} B_{\nu} - \partial_{\nu} B_{\mu} \\ W^a_{\mu\nu} &= \partial_{\mu} W^a_{\nu} - \partial_{\nu} W^a_{\mu} - g f_{abc} W^b_{\mu} W^c_{\nu}, \end{aligned} \tag{1.8}$$

[2] In some literature, the g is also quoted as g_W, where W denotes the weak interaction. We use g throughout this chapter.

where f_{abc} are the structure constants of the algebra of $SU(2)_L$; parameter $Y = 2(Q - I_W^3)$ is the weak hypercharge, which combines the electromagnetic charge Q and the third component of weak isospin I_W^3. The $T_a(a = 1, 2, 3)$ are the generators of $SU(2)$, which can be represented by the Pauli matrices $\frac{\sigma_a}{2}$.

The physical gauge bosons can be obtained by the linear combination of the gauge fields. The W bosons, $W_\mu^\pm$, are straightforwardly determined by the charged-current interaction

$$W_\mu^+ = \frac{1}{\sqrt{2}}(W_\mu^1 - iW_\mu^2) \qquad (1.9)$$
$$W_\mu^- = \frac{1}{\sqrt{2}}(W_\mu^1 + iW_\mu^2).$$

For the photon and the Z boson, A_μ and Z_μ, the linear combination is an outcome of the Higgs mechanism (see Section 1.1.2).

$$A_\mu = +B_\mu \cos\theta_W + W_\mu^3 \sin\theta_W \qquad (1.10)$$
$$Z_\mu = -B_\mu \sin\theta_W + W_\mu^3 \cos\theta_W.$$

Here, θ_W is the weak mixing angle ($\sin^2\theta_W \simeq 0.23$). It relates to the electroweak interaction constants

$$g \sin\theta_W = g' \cos\theta_W. \qquad (1.11)$$

1.1.2 Spontaneously electroweak symmetry breaking

The Lagrangians shown in the previous section describe massless gauge bosons and fermions. In fact, adding mass terms such as $\frac{1}{2}m_\gamma^2 A_\mu A^\mu$ and $-m_F \bar{F}F$ violates the local gauge symmetry. In the SM, the Higgs mechanism [17–22] spontaneously breaks the electroweak $SU(2)_L \times U(1)_Y$ local gauge symmetry, generates mass to particles, and predicts a physical massive scalar particle.

The SM chooses the most economic form for the Higgs field — a weak isospin doublet (Higgs doublet) with weak hypercharge $Y = 1$. The Higgs doublet consists of four real scalar fields

$$\phi = \begin{pmatrix} \phi^+ \\ \phi^0 \end{pmatrix} = \frac{1}{\sqrt{2}} \begin{pmatrix} \phi_1 + i\phi_2 \\ \phi_3 + i\phi_4 \end{pmatrix}. \qquad (1.12)$$

Therefore the gauge invariant Lagrangian of the Higgs field can be termed as

$$\mathscr{L}_{\text{Higgs}} = (D_\mu \phi)^\dagger (D_\mu \phi) - V(\phi), \qquad (1.13)$$

with the covariant derivative

$$D_\mu = \partial_\mu + ig' \frac{Y}{2} B_\mu + igT_a W_\mu^a, \qquad (1.14)$$

and the Higgs potential

$$V(\phi) = \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2. \tag{1.15}$$

In the Higgs potential, the coefficients satisfy the condition $\mu^2 < 0$ and $\lambda > 0$. This indicates degenerated minima of the Higgs potential at

$$\phi^\dagger\phi = \frac{v^2}{2} = -\frac{\mu^2}{2\lambda}. \tag{1.16}$$

The minima are called the vacuum. An one-dimensional illustration of the Higgs potential is shown in Figure 1.1.

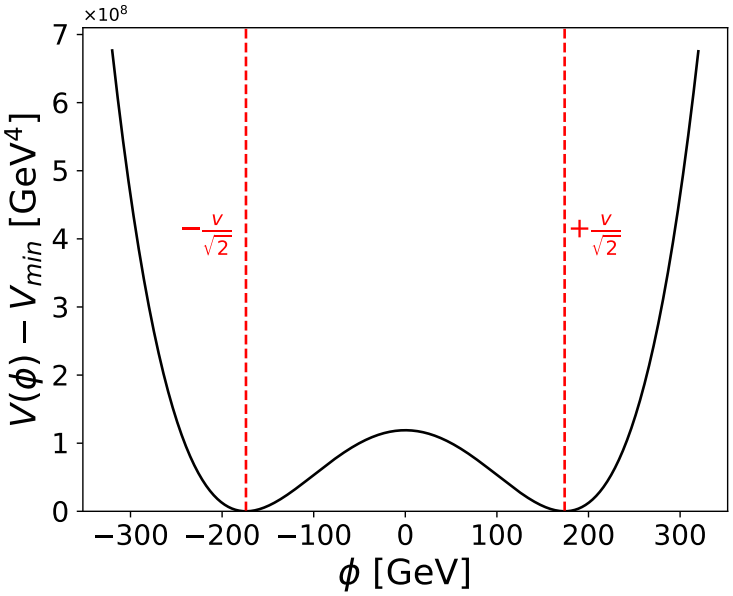


Figure 1.1: One-dimensional illustration of the Higgs potential $V(\phi) = \mu^2\phi^2 + \lambda\phi^4$, with $\mu^2 < 0$ and $\lambda > 0$. The degenerated minima are at $\pm\frac{v}{\sqrt{2}}$. The constant minimum value is subtracted from the potential.

The $U(1)$ gauge symmetry of the electromagnetic interaction must remain unbroken for the photon to be massless. Therefore the choice of the vacuum expectation value (VEV) of Higgs doublet is $\phi_0 = \frac{1}{\sqrt{2}}\binom{0}{v}$. The value of v is fixed by the Fermi coupling strength G_F ($v = (\sqrt{2}G_F)^{-1/2} = 246$ GeV). Due to gauge invariance, the expansion from the vacuum can be written in unitary gauge

$$\phi(x) = \frac{1}{\sqrt{2}}\begin{pmatrix} 0 \\ v + h(x) \end{pmatrix}, \tag{1.17}$$

where $h(x)$ is the Higgs field. By substituting it into the Higgs Lagrangian, the massive scalar field h, the terms related to the masses of electroweak gauge fields, the coupling of Higgs field to gauge

bosons fields (see Section 1.2.1) and the self couplings of h are derived

$$\begin{aligned}\mathscr{L}_{\text{Higgs}} = & \underbrace{\frac{1}{2}(\partial_\mu h)(\partial^\mu h) - \lambda v^2 h^2}_{\text{massive scalar field } h} - \underbrace{\lambda v h^3 - \frac{\lambda}{4}h^4}_{h \text{ self-interaction}} \\ & + \underbrace{\frac{v^2}{8}g^2(W_\mu^1 W^{1,\mu} + W_\mu^2 i W^{2,\mu}) + \frac{v^2}{8}(gW_\mu^3 - g'B_\mu)(gW^{3,\mu} - g'B^\mu)}_{\text{terms that generate the mass of gauge bosons}} \\ & + \underbrace{\mathscr{L}_{\text{h, gauge}}}_{h\text{, gauge boson interaction}} .\end{aligned} \tag{1.18}$$

The mass of the physical scalar field h (the Higgs boson) is

$$m_h = \sqrt{2\lambda}v. \tag{1.19}$$

Therefore the parameter λ that relates to the self-interaction of Higgs field can be determined by measuring the mass of the Higgs boson, which is $\lambda = 0.13$. The self interaction of Higgs boson is illustrated in Figure 1.2, where the coupling strengths are labelled near the Feynman vertex.

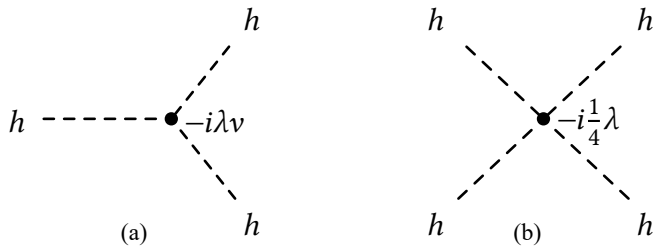


Figure 1.2: Feynman rules for the self interaction of Higgs boson. (a) Trilinear (b) Quartic.

The masses of the physical gauge bosons can be obtained from the second row of Equation (1.18). The mass of the $W^{1,2}$ fields can be read from the Lagrangian directly. It is the same for the physical $W^\pm$ bosons. The masses of photon and Z boson can be derived by diagonalising the mass matrix of the quadratic term of B and W^3. Concurrently, the linear combination aforementioned in Equation (1.10) is obtained. The masses of physical gauge bosons are

$$m_{W^\pm} = \frac{1}{2}gv \qquad m_\gamma = 0 \qquad m_Z = \frac{1}{2}v\sqrt{g^2 + g'^2} = \frac{m_{W^\pm}}{\cos\theta_W}. \tag{1.20}$$

The Lagrangian of interactions between h and gauge fields $\mathscr{L}_{\text{h, gauge}}$ contains terms in form of hVV and $hhVV$, where V denotes a vector boson field. These couplings are shown in Figure 1.3.

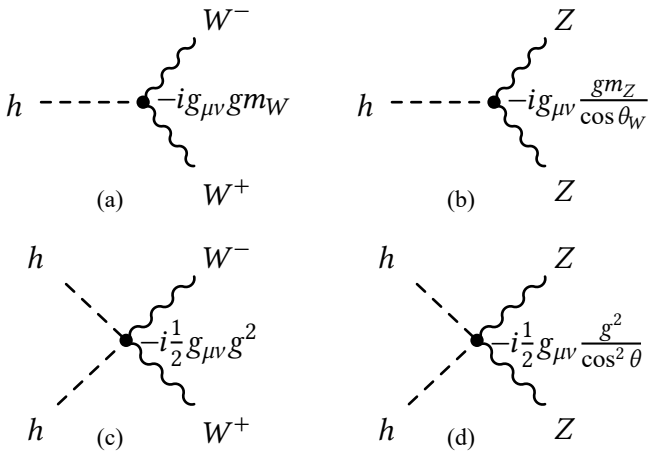


Figure 1.3: Feynman rules for the interaction between Higgs boson and $W^{\pm}/Z$ boson. (a) hW^+W^-, (b) hZZ, (c) hhW^+W^-, (d) $hhZZ$.

1.1.2.1 The mass of the fermions

The mass of fermions are also generated via Higgs mechanism. The Higgs doublet is embedded in the Dirac Lagrangian while retaining the $SU(2)_L \times U(1)_Y$ gauge invariance of the Lagrangian. It is

$$\mathscr{L}_{\text{Fermion}} = \underbrace{-G_D(\bar{L}\phi R + \bar{R}\phi^\dagger L)}_{\text{for down type fermions}} - \underbrace{G_U(\bar{L}\phi_c R + \bar{R}\phi_c^\dagger L)}_{\text{for upper type fermions}}, \tag{1.21}$$

where $L(x)$ and $R(x)$ are the weak isospin doublet and singlet fermion fields, $\phi_c = -i\sigma_2\phi^*$ is the weak hypercharge conjugate of the Higgs double ϕ, G_D and G_U are fermion flavour dependent coupling constants for down type and upper type fermions, respectively. By substituting the vacuum expansion of Higgs doublet, the Dirac Lagrangian of one of the fermions (denoted as F) becomes

$$\mathscr{L}_{\text{Fermion}} = \underbrace{-\frac{G}{\sqrt{2}}v(\bar{F}_L F_R + \bar{F}_R F_L)}_{\text{fermion mass}} - \underbrace{\frac{G}{\sqrt{2}}h(\bar{F}_L F_R + \bar{F}_R F_L)}_{h\text{, fermion Yukawa coupling}} . \tag{1.22}$$

This Lagrangian consists of the following two terms. The fermion mass term where $m_F = \frac{1}{\sqrt{2}}Gv$, and the coupling term that describes the coupling between Higgs boson and the fermion. The coupling is called the Yukawa coupling, as shown in Figure 1.4. The Yukawa coupling constant G's are determined by the observed fermion masses.

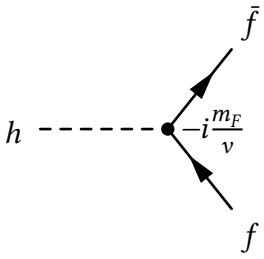


Figure 1.4: Feynman rule for the coupling between Higgs boson and fermion.

1.2 Higgs boson and pair production

1.2.1 Properties of Higgs boson

In 2012, ATLAS [1] and CMS [2] experiment discovered a resonance particle with a mass of approximately 125 GeV. It is the only scalar particle ever observed, whose properties are highly consistent with those of the Higgs boson predicted by the electroweak symmetry breaking mechanism. Thus, this particle is known as the SM Higgs boson (H) [3]. The experimental value of the mass of the Higgs boson is $m_H = 125.10 \pm 0.14$ GeV and the spin-parity is $J^P = 0^+$ [23].

At the LHC, the Higgs boson is dominantly produced via the following processes: gluon-gluon fusion via a loop of t/b-quark (ggF); vector boson fusion (VBF); vector boson associated process (VH, i.e. WH and ZH); top-quark pair associated process (ttH). The Leading Order (LO) Feynman diagrams of these production modes are summarised in Figure 1.5 with the theoretical prediction of the cross sections with a centre-of-mass (COM) energy of $\sqrt{s} = 13$ TeV and Higgs boson mass of 125 GeV. The cross sections are calculated with state-of-art theoretical prediction in the high order perturbation calculation of QCD and EW theory [5].

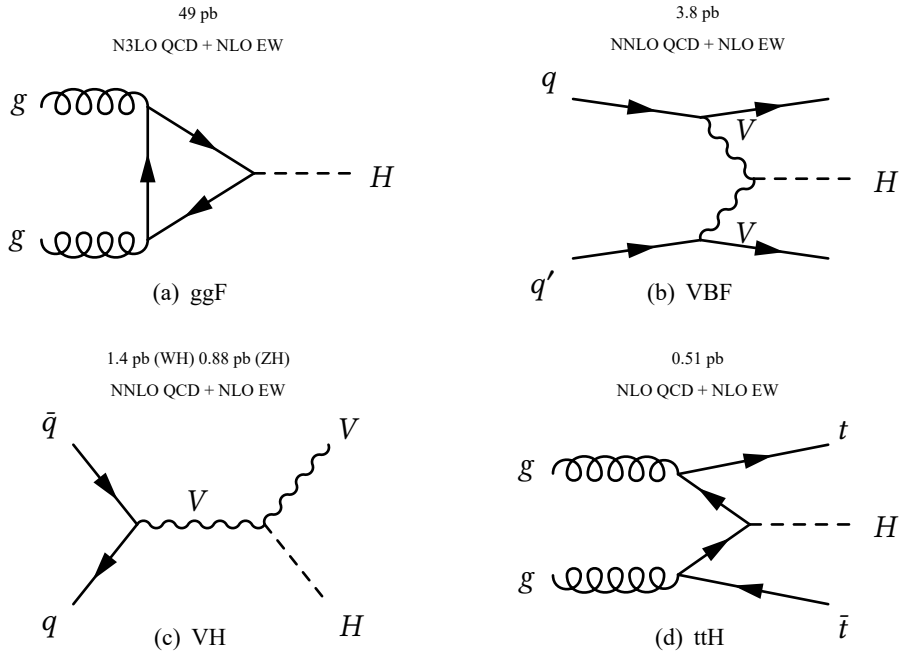


Figure 1.5: The leading order Feynman diagrams of the major production modes of Higgs boson via p-p collision at the LHC. The corresponding cross sections with high order QCD and EW theory corrections are shown at the top of each sub-figures [5]. 'N' denotes 'next-to'.

As implied by Section 1.1.2, the SM Higgs mechanism fixes the couplings of the Higgs boson to vector bosons and fermions, and the coupling strengths are proportional to the mass of the particles. Based on that, the decay mechanism of Higgs boson can be predicted. Note that one of the vector

[3] To clarify the notation. Last chapter uses lowercase h to denote the Higgs field as well as the physical particle Higgs boson. From this chapter on, the capital H refers specifically to the physical Higgs boson.

boson must be off-shell in $H \to VV$ due to the mass threshold, and the Higgs boson only effectively couples to massless photon and gluons via a loop diagram [4]. Table 1.1 summarises the theoretical prediction of the branching ratios of major Higgs boson decay channels, assuming $m_H = 125$ GeV.

Table 1.1: The branching ratios of major Higgs boson decay channels assuming $m_H = 125$ GeV [23].

Decay channel	Branching ratio
$H \to b\bar{b}$	$58.2^{+1.2}_{-1.3}$ %
$H \to W^+W^-$	21.4 ±1.5 %
$H \to \tau^+\tau^-$	6.27 ±1.6 %
$H \to ZZ$	2.62 ±1.5 %
$H \to \gamma\gamma$	0.227 ±2.1 %

The mass of the Higgs boson is a free parameter in the SM. It also determines the value of Higgs boson self-coupling constant λ. Direct measurement on λ, is one of the desire to search for Higgs boson pair (HH) production because varying λ can significantly change the production cross section of HH (denoted as σ_{HH}). The relation between σ_{HH} and λ is explained in Section 1.2.2. Conventionally, the measurement is usually on the coupling modifier $\kappa_\lambda = \frac{\lambda}{\lambda_{SM}}$, where $\lambda_{SM} \simeq 0.13$ is the SM value of λ. A constraint on κ_λ is set using the $HH \to b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ decay channel in Section 4.8. In the next subsection, a detailed description of the production and decay mechanism of Higgs boson pair is given.

1.2.2 Production and decay of Higgs boson pair

Similar as the production of a single Higgs boson, the dominant production modes of Higgs boson pair are gluon-gluon fusion and vector boson fusion. The LO Feynman diagrams are summarised in Figure 1.6. The labels for the Feynman diagrams are the same as Figure 1.5 for the single Higgs boson case. They are not displayed here for clarity.

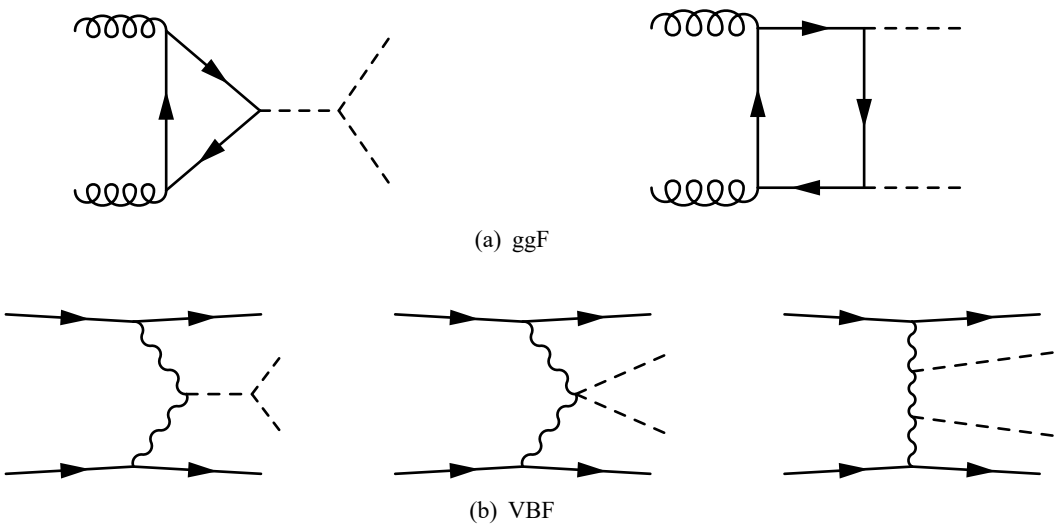
The state-of-art theoretical prediction of the ggF cross section is calculated at Next-Next-to LO (NNLO) precision [5] [5] using FTapprox [24] method, whose uncertainty is given by Ref. [25]. The cross section is

$$\sigma_{\text{ggF}} = 31.05 \text{ fb} \pm 3.0\%(\text{PDF and } \alpha_S)^{+2.2\%}_{-5.0\%}(\text{Scale})^{+4\%}_{-18\%}(m_{\text{top}}). \tag{1.23}$$

In the FTApprox method, real radiation corrections are calculated with finite top mass, while the virtual loop corrections are based on Higgs Effective Field Theory (HEFT). The uncertainties are from the parton distribution function and α_S of proton based on PDF4LHCNNLOMC, the scale of renormalisation and factorisation, and the choice of renormalisation scheme and scale in the cal-

[4] See Figure 2.4 and 2.7 of Ref. [13]

[5] In this section, the precisions are QCD precision if no specification. Electroweak calculation is always at NLO precision.



(a) ggF

(b) VBF

Figure 1.6: The leading order Feynman diagrams of ggF (a) and VBF (b) production modes of Higgs boson pair via p-p collision at the LHC.

culation of top-quark mass. For VBF, the cross section is calculated at Next-Next-Next-to LO (N3LO) [26],

$$\sigma_{\text{VBF}} = 1.726 \text{ fb} \pm 2.1\%(\text{PDF and } \alpha_S)^{+0.03\%}_{-0.04\%}(\text{Scale}). \tag{1.24}$$

The cross section of Higgs boson pair production is about one thousandth of that of the single Higgs boson, because of the destructive interference between the 'triangle' and 'box' diagrams in Figure 1.6(a). The suppressed cross section makes the search very challenging. The production cross section can be modified by varying the Higgs boson self-coupling constant. Figure 1.7 shows the production cross section as a function of κ_λ.

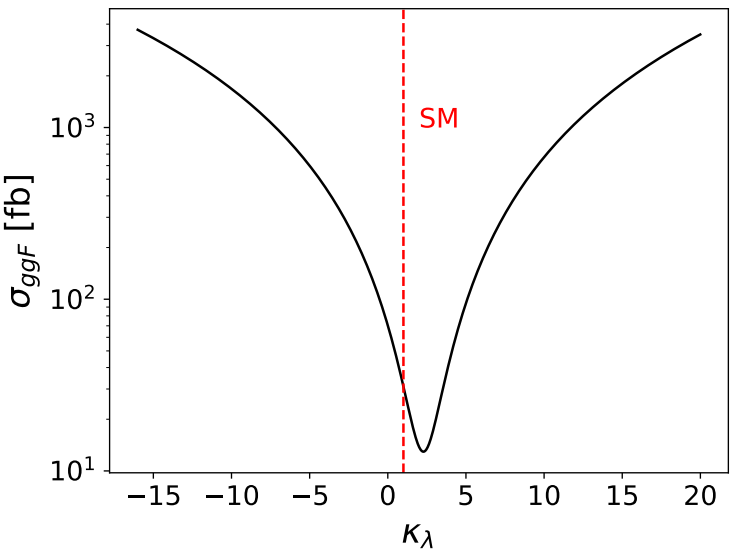


Figure 1.7: The production cross section of Higgs boson pairs via gluon-gluon fusion as a function of κ_λ, calculated based on Ref. [5]. The red vertical line represents the SM case $\kappa_\lambda = 1$.

The cross sections are used to normalise the simulated HH samples used in Chapter 4, which

are generated by POWHEG BOX V2 [27] with full top-quark mass (FT) at Next-to LO (NLO) precision. In the FT method, finite mass of top-quark is used for both real and virtual corrections. The FT method is used to generate the HH samples because it provides better predictions of Higgs boson kinematics [28]. Comparisons between the kinematics of the HH system in this sample and that in the FTApprox-based sample generated using MADGRAPH5_aMC@NLO [29, 30] are shown in Figure 1.8. This results are reported in Ref. [31]. They are important materials for the decision of the ggF sample generators for HH physics analysis in ATLAS.

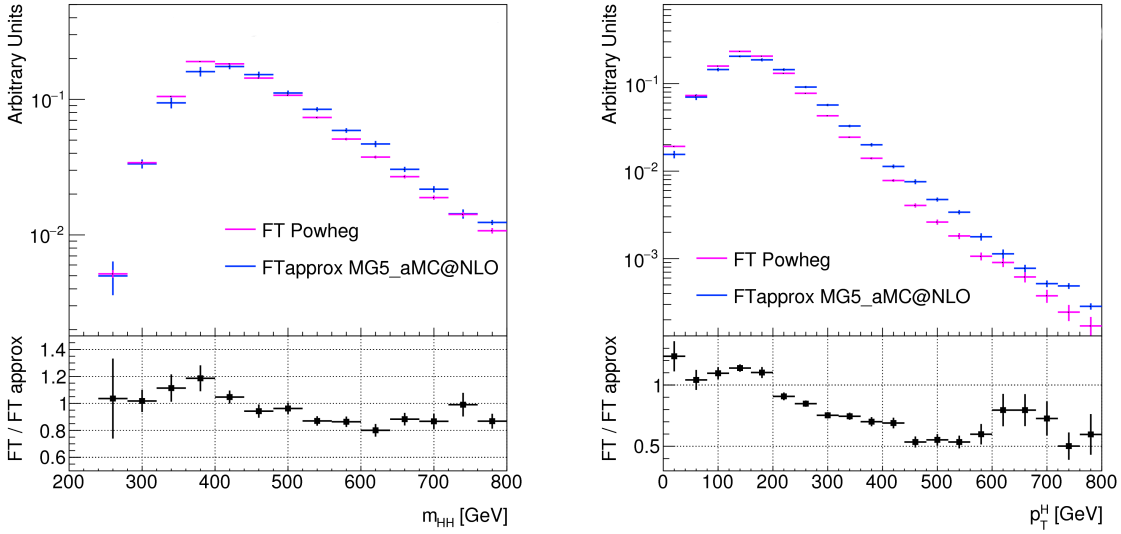


Figure 1.8: Distribution of invariant mass of HH system m_{HH} (left) and the transverse momentum of the leading Higgs boson p_T^H (right) of the ggF production at $\sqrt{s} = 13$ TeV. The samples generated using FT by POWHEG BOX V2 and those generated using FTApprox by MADGRAPH5_aMC@NLO are compared.

The decay of a pair of Higgs boson is essentially the combination of the decays of each Higgs boson. Therefore, the branching ratios of the main decay channels of HH can be calculated based on Table 1.1. They are summarised in Table 1.2.

Table 1.2: The branching ratios of the major Higgs boson pair decay channels, assuming $m_H = 125$ GeV

Decay channel	Branching ratio
$HH \to b\bar{b}b\bar{b}$	33.8%
$HH \to b\bar{b}W^+W^-$	24.9%
$HH \to W^+W^-W^+W^-$	4.58%
$HH \to b\bar{b}\tau^+\tau^-$	7.30%
$HH \to b\bar{b}\gamma\gamma$	0.264%

1.2.3 Global properties of the Higgs potential

Measuring the self-coupling of the Higgs boson is of great importance in the understanding of the electroweak symmetry breaking mechanism. More intriguingly, the self-coupling reveals the global information of the Higgs potential. The information is crucial to profound physics problems. Two interesting topics intimately related to the global properties of the Higgs potential are briefly introduced below. The main goal of the search for Higgs pair production in this thesis is to probe any processes that might affect the global properties of the Higgs potential.

1.2.3.1 Stability of vacuum

The shape of the Higgs potential determines the phase of the SM (see Figure 1.9). At EW scale, the tree level potential described in Section 1.1.2 is a good approximation. However, as the value of the Higgs field h moving towards the higher energy, the quantum correction must be taken into account, by assuming that the SM is valid up to the Planck scale $\Lambda_{\text{Pl}} \sim 10^{19}$ GeV.

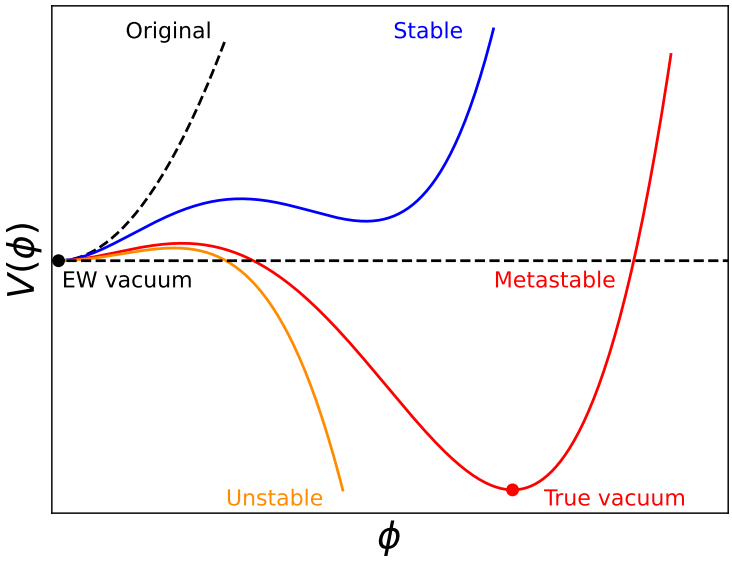


Figure 1.9: Schematic illustration of the effective Higgs potential in different phases of SM. The value of h starts from the EW vacuum of the SM $h = 0$. The black dashed line represents the tree level potential, shown as a reference. The blue, red and orange solid lines represent the potentials whose shapes correspond to the stable, metastable and unstable phases of SM, respectively. A true vacuum at high value of h exists in the metastable phase. The scale in the figure does not correspond to the real scale of h or $V(h)$.

The SM parameters are calculated at NNLO precision at the EW scale and extrapolated to the Planck scale using Renormalisation Group Equations (RGE) in Ref. [32]. Due to the evolution of the parameters from low to high energy scale, the shape of the Higgs potential changes vitally. It is found that the Higgs quartic coupling constant and the top-quark Yukawa coupling constant is very sensitive to the energy scale (e.g. Figure 1 of Ref. [33]). In fact, the current measurements on the masses of the Higgs boson and the top-quark show that the phase of the SM is near the

boundary of stable and metastable phase and is compatible with a metastable vacuum [32]. As a consequence of that, there is a probability for the SM to transition to the true vacuum by quantum tunnelling effect. This is particularly interesting from the cosmological point of view, because our universe would be a completely different place if the tunnelling rate is sufficiently high to induce a phase transition. As discussed in Ref. [34], the decay rate of our EW vacuum is enormously small. Thus, we're safe unless the decay rate is significantly changed by new physics mechanism.

1.2.3.2 Electroweak baryongenesis

The evolution of the Higgs potential in the thermal history of early universe can be simulated by the finite temperature quantum field theory. It implies the dynamics of the electroweak phase transition (EWPT), by which the EW theory transforms from the symmetric phase to the symmetry breaking phase. In particular, the evolution reveals the type of the EWPT — first-order or second-order phase transition.

First-order EWPT is required by the electroweak baryongenesis (EWBG) [35], which is one of the most promising theory that may explain the observed matter-antimatter asymmetry in the universe. Figure 1.10 schematically demonstrates how the form of the Higgs potential changes as the universe cools down, assuming that the EWPT is first-order phase transition. When the temperature T is very high, only phase with zero VEV exists. Once the temperature reaches a critical value T_c, a degenerated vacuum with VEV=ϕ_c is created. As the universe further cools down to $T \to 0$, the symmetric phase tunnels to the symmetry-breaking phase. Finally, the minimum field value converges to the EW vacuum v.

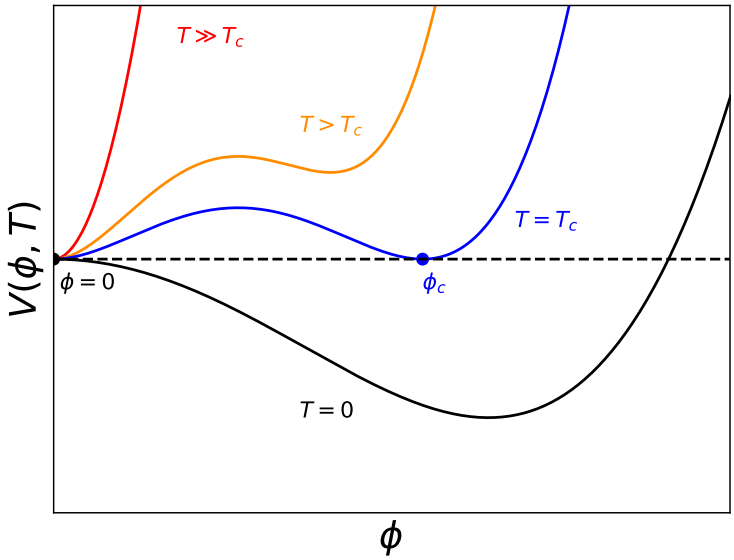


Figure 1.10: Schematic illustration of the effective Higgs potential under different thermal scenarios of the universe, assuming the first-order electroweak phase transition. The scale in the figure does not correspond to the real scale of h or $V(h)$.

Furthermore, the phase transition needs to be strong (so-called strong first-order phase transi-

tion) otherwise the baryon numbers are washed out by sphalerons. The necessary condition is

$$\frac{\phi_c}{T_c} \gtrsim 1. \tag{1.25}$$

However, the measurements on the Higgs boson indicate that the EWPT under the SM is a second-order phase transition (a.k.a crossover) [36, 37], by which the VEV smoothly transition from 0 to v. Thus, new physics beyond the SM are needed to produce the desired strong first-order phase transition. For instance, it is predicted by Ref. [38] that if there is an enhancement of more than 50% to the strength of SM Higgs boson self-coupling, i.e. $\kappa_\lambda \gtrsim 1.5$, then the EWPT can become strong first-order phase transition.

1.2.4 Beyond-the-SM physics and Higgs pair production

Whilst aspiring to probe the global properties of the Higgs potential, the actual approach of the collider is to search for the production of Higgs boson pairs. This is because modifications to the Higgs potential can change the cross section of Higgs boson pair production. And conversely, if deviations on the cross section of Higgs boson pair production are observed against the SM prediction, it will renew our understanding of the nature of Higgs potential.

In this thesis, both the resonant and non-resonant HH production mode are searched. The searches are conducted in a model-independent way, meaning that they are sensitive to any new HH production signals. The baseline signal model in the search for resonant HH production is a heavy, narrow-width, scalar particle X with a mass of $m_X > 2m_H$ that decays to a pair of SM Higgs bosons, as illustrated in Figure 1.11), while that for the non-resonant HH production is the SM ggF+VBF HH production shown in Figure 1.6.

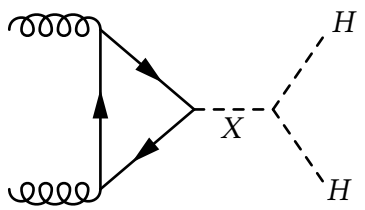


Figure 1.11: Feynman diagrams of the resonant production of Higgs boson pairs via the decay of a scalar particle considered in the thesis.

1.2.4.1 Resonant production

New heavy resonant particles that decays to a pair of Higgs boson are predicted by various BSM models, a review can be found in Ref. [6]. For instance, a real scalar singlet ξ can be added in a

simple extension of the SM [39–41]. The extended Higgs potential can be written as

$$\begin{aligned} V(\phi,\xi) = {} & \mu^2\phi^\dagger\phi + \lambda(\phi^\dagger\phi)^2 \\ & + \frac{a_1}{2}\phi^\dagger\phi\xi + \frac{a_2}{2}\phi^\dagger\phi\xi^2 \\ & + b_1\xi + \frac{b_2}{2}\xi^2 + \frac{b_3}{3}\xi^3 + \frac{b_4}{4}\xi^4. \end{aligned} \tag{1.26}$$

After the spontaneous symmetry breaking, by taking the unitary gauge $\xi = \frac{1}{\sqrt{2}}(v_\xi + s)$ and $\phi = \frac{1}{\sqrt{2}}\binom{0}{v+h}$, the h and s fields are mixed in the form

$$\begin{pmatrix} h_1 \\ h_2 \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} \begin{pmatrix} h \\ s \end{pmatrix}, \tag{1.27}$$

where v_ξ is the VEV of ξ and s is a new physical scalar field, and h_1, h_2 are the mass eigenstates with eigenvalues m_1, m_2. If h_1 is the SM Higgs boson ($m_1 = m_H$) and the mass of h_2 satisfies $m_2 > 2m_1$, then the resonant production of HH via the decay of $h_2 \to h_1h_1$ is allowed. Attractively, with the help of the new scalar, the new Higgs potential in its effective form under the finite temperature QFT can provide the conditions for a strong first-order EWPT, while giving compatible predictions on the Higgs boson with the SM. This feature is a strong motivation to search for resonant Higgs pair production.

1.2.4.2 Non-resonant production

Various BSM models predict enhancement of Higgs pair production rate via non-resonant processes, either by new particles that contribute to the loop diagram in the ggF production mode [6], or by anomalous coupling, e.g. non-SM Higgs boson self-coupling or non-SM top-quark Yukawa coupling [42, 43]. We consider the non-SM Higgs boson self-coupling coupling parameterized by the modifier κ_λ in the ggF HH production mode. In this case, the variation of κ_λ linearly alters the amplitude of the 'triangle' diagram in Figure 1.6(a). Thus, the total amplitude of the ggF HH production can be written as

$$\mathscr{A}(\kappa_\lambda) = \mathscr{A}_\square + \kappa_\lambda \mathscr{A}_\triangle, \tag{1.28}$$

where $\mathscr{A}_\square$ and $\mathscr{A}_\triangle$ are the amplitude of 'box' and 'triangle' diagrams. The production cross section σ_{ggF} is proportional to

$$\begin{aligned} \sigma_{\text{ggF}} \propto |\mathscr{A}(\kappa_\lambda)|^2 &= |\mathscr{A}_\square + \kappa_\lambda \mathscr{A}_\triangle|^2 \\ &= \underbrace{|\mathscr{A}_\square|^2}_{\text{'box' term}} + \kappa_\lambda \underbrace{(\mathscr{A}_\square^* \mathscr{A}_\triangle + \mathscr{A}_\triangle^* \mathscr{A}_\square)}_{\text{interference term}} + \kappa_\lambda^2 \underbrace{|\mathscr{A}_\triangle|^2}_{\text{'triangle' term}} \end{aligned} \tag{1.29}$$

The total cross section σ_{ggF} as a function of κ_λ has been shown in 1.7. In addition, it is found that the differential cross section over the kinematics of the Higgs boson (pair) is also affected by

the variation of κ_λ. As shown in Figure 1.12, when $\kappa_\lambda = 0$, i.e. the scenario with no Higgs boson self-coupling, the HH invariant mass m_{HH} mostly distributes above the double top-quark mass $2m_t \simeq 346$ GeV. When $\kappa_\lambda = 5$, the self-coupling is enhanced and the 'triangle' diagram dominates the HH production. In this case, the m_{HH} mostly distributes at the threshold of double Higgs boson mass $2m_H \simeq 250$ GeV. The scenario of $\kappa_\lambda = 2.4$ corresponds to the case of maximum interference between 'box' and 'triangle' diagrams. A dip is produced between $2m_H$ and $2m_t$.

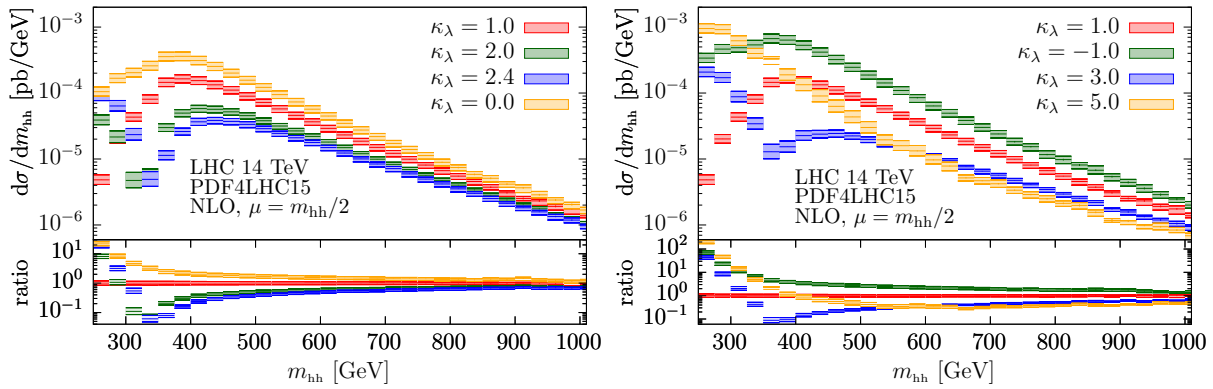


Figure 1.12: Distributions of the invariant mass of Higgs boson pair at various value of κ_λ [44].

To set constraints on the value of κ_λ, its value needs to be finely scanned. Therefore the simulated HH samples with arbitrary κ_λ values are required. Practically, samples with specific □values are represented by a linear combination of three basis samples. An empirical choice of the basis samples is three samples generated with $\kappa_\lambda = 0, 1, 20$. The coefficients are derived by substituting the 'box', 'triangle' and interference terms in Equation (1.29) with the basis amplitudes. The final linear combination is

$$\begin{aligned} |\mathscr{A}(\kappa_\lambda)|^2 = &(1 - \frac{399}{380}\kappa_\lambda + \frac{1}{20}\kappa_\lambda^2)|\mathscr{A}(0)|^2 \\ &+ (\frac{20}{19}\kappa_\lambda - \frac{1}{19}\kappa_\lambda^2)|\mathscr{A}(1)|^2 \\ &+ (-\frac{1}{380}\kappa_\lambda + \frac{1}{380}\kappa_\lambda^2)|\mathscr{A}(20)|^2. \end{aligned} \tag{1.30}$$

Chapter 2

The ATLAS Experiment

This chapter provides a brief introduction to the foundations of performing physics analyses in the ATLAS experiment at the LHC. The concepts in the LHC and the ATLAS experiment are from Ref. [45] and [46].

2.1 The Large Hadron Collider

The Large Hadron Collider (LHC) [47–49] is a platform of advanced hadron accelerator and collider. The prime usage is to collide protons with a designed maximum centre-of-mass energy of $\sqrt{s} = 14$ TeV and peak luminosity of $\mathscr{L} = 10^{34}$ cm^{-2}s^{-1}, and it also runs with heavy-ion beams. The LHC ring is 26.7 kilometres long, installed between 45 and 170 meters underground. The ring is linked to the CERN accelerator complex by two 2.5-kilometre-long transfer tunnels. The hadron beams are injected in the ring through the tunnels.

The hadrons are bunched together while being transferred. In proton runs, there are 2808 bunches in the ring. Each bunch contains about 10^8 protons. The interval between the bunches is typically 25 nanoseconds, corresponding to a bunch crossing frequency of 40 MHz.

The injected beams are captured, accelerated and stored in superconducting radio frequency cavities. Two beams travel with almost the speed of light in opposite direction and in separate cavities. They collide at four interaction areas of the LHC ring, supporting four largest experiments and several smaller ones. Among the four major experiments, the ATLAS [50] and the CMS [51] experiments are general-purpose experiments that are designed to explore SM, Higgs boson, supersymmetry and various exotic physics, while the LHCb [52] experiment is designed to study B-physics and ALICE [53] experiment is dedicated to study quark-gluon plasma using ion collisions.

Figure 2.1 outlines the accelerator complex and the experiment sites of the LHC. In the case of proton beams, the injection chain is *Linac→ Proton Synchrotron Booster (Booster) → CERN Proton Synchrotron (PS)→ Super Proton Synchrotron (SPS)*. The flow of the proton is labelled by *p* and grey-blue arrow in the figure.

The operation of the LHC is scheduled in periods. In the first period in year of 2011 and 2012,

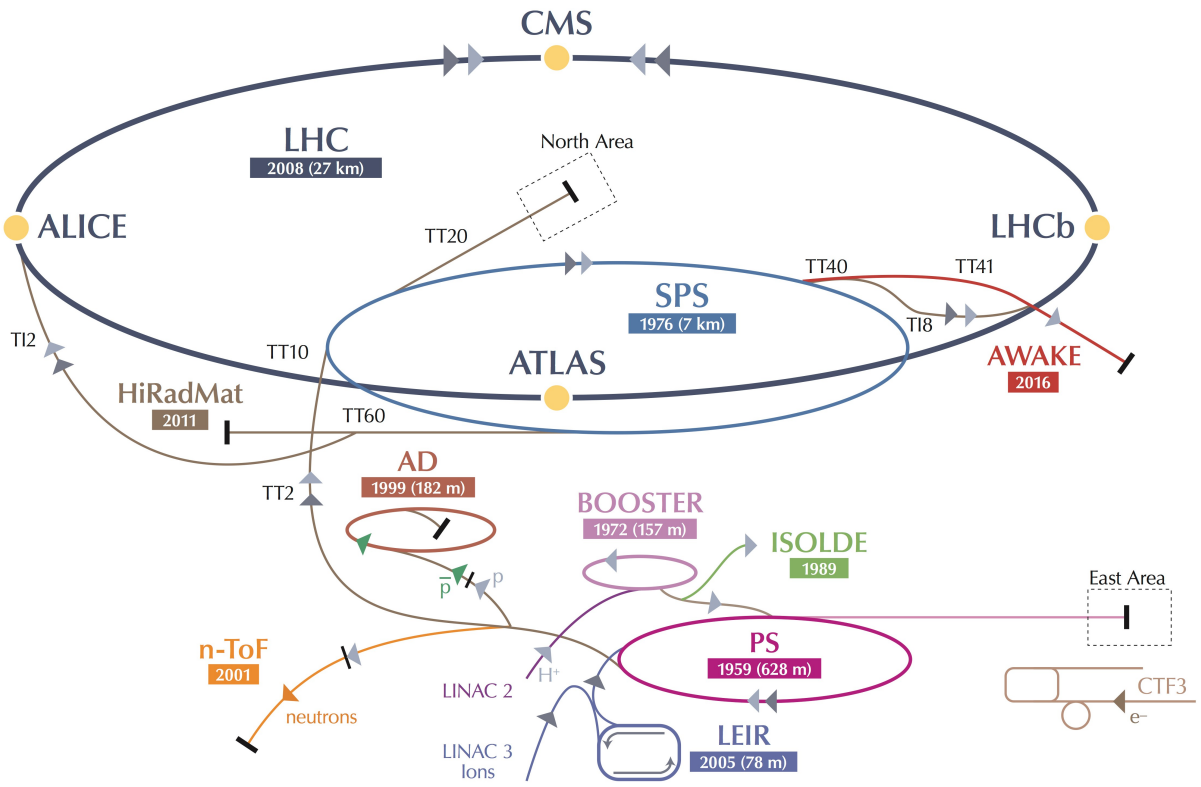


Figure 2.1: The accelerator complex and the experiment sites of the LHC [54].

namely Run-1, the protons were collided at COM energy $\sqrt{s} = 7$ and 8 TeV, delivering p-p collision dataset of integrated luminosity L =28.3 fb^{-1} to the ATLAS detector. The discovery of the Higgs boson took place in this period. Data-taking from 2015 to 2018 is the Run-2 period of LHC. The instantaneous luminosity was increased, especially for the runs in 2017 and 2018, and the COM energy was upgraded to 13 TeV. Finally, LHC delivered L =156 fb^{-1} p-p collisions to the ATLAS detector, from which 139 fb^{-1} are qualified for physics analysis. The cumulation of luminosity and the average number of interactions per proton bunch crossing $\langle\mu\rangle$ (so-called pile-up) in Run-2 of LHC are summarised in Figure 2.2. The majority of the data are taken in 2017 and 2018, during which $\langle\mu\rangle$ is also higher than the previous years.

At the moment, the LHC is during its second Long Shutdown (LS2) and is under preparation for the Run-3 period. In Run-3, the LHC is expected to run with a setup of $\sqrt{s} = 13.5$ TeV or 14 TeV and $\mathscr{L} = 2\times10^{34}$ cm^{-2}s^{-1}, and to accumulate 300 fb^{-1} p-p collision data. After that, the LHC will be upgraded into the High-Luminosity LHC (HL-LHC) [56], where the instantaneous luminosity will be increased to more than 5×10^{34} cm^{-2}s^{-1}. A dataset of an integrated luminosity of 3000 fb^{-1} is expected to be delivered.

2.2 The ATLAS detector

The ATLAS (A Toroidal LHC ApparatuS) detector [50] locates at one of the interaction points of the LHC. This scientific machine is surprisingly large – 25 meters in height, 44 meters in length and

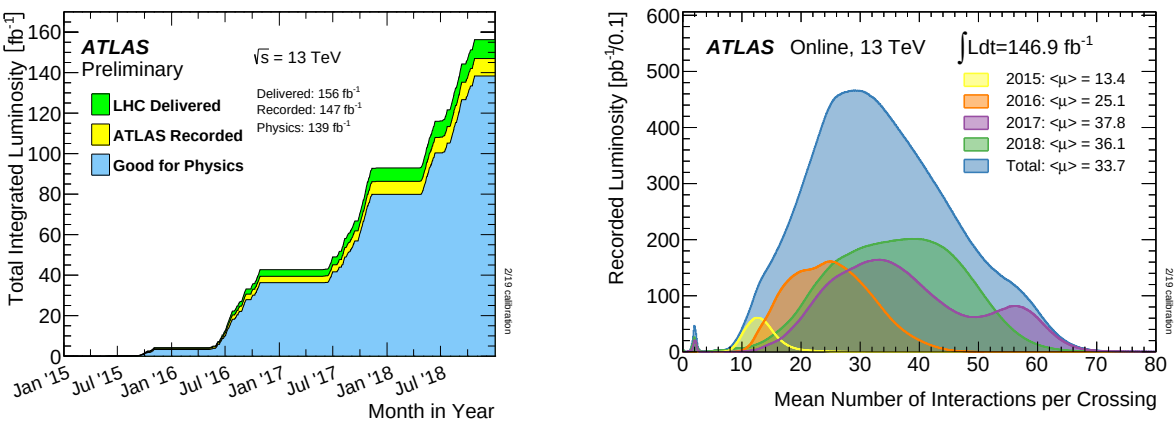


Figure 2.2: Left: The cumulation of luminosity as a function of date during the Run2 p-p collision data-taking. The green, yellow and blue curves correspond to the data that were delivered by LHC, recorded by ATLAS detector, and pass the good quality criteria by ATLAS experiment. Right: The mean number of interactions per proton bunch crossing of the ATLAS recorded dataset [55].

7000 tons in weight. As demonstrated in Figure 2.3, from near the beamline to the outermost layer, the particle detectors include the inner detector, the electromagnetic and hadronic calorimeters and the muon spectrometer.

The ATLAS detector measures the energy and direction of the outgoing particles inside very dense fluxes of particles produced by a bunch crossing. At the same time, it provides excellent identification and energy resolution for all kinds of particles (electrons, muons, photons, pions, neutrons, protons). An overview of the targeted performance of each component of the ATLAS detector is shown in Table 2.1.

Table 2.1: An overview of the targeted performance of the components of the ATLAS detector [46].

Component	Resolution (p_{T} and E in GeV)	η coverage	
		Measurement	Trigger
Inner Detector	$\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 0.05\% \cdot p_{\mathrm{T}} \oplus 1\%$	$\lvert\eta\rvert < 2.5$	
Electromagnetic calorimeter	$\sigma_E/E = 10\%/\sqrt{E} \oplus 0.7\%$	$\lvert\eta\rvert < 3.2$	$\lvert\eta\rvert < 2.5$
Hadronic calorimeter			
Barrel and end-cap	$\sigma_E/E = 50\%/\sqrt{E} \oplus 3\%$	$\lvert\eta\rvert < 3.2$	$\lvert\eta\rvert < 2.5$
Forward	$\sigma_E/E = 100\%/\sqrt{E} \oplus 10\%$	$3.1 < \lvert\eta\rvert < 4.9$	$3.1 < \lvert\eta\rvert < 4.9$
Muon spectrometer	$\sigma_{p_{\mathrm{T}}}/p_{\mathrm{T}} = 10\%$ at p_{T} =1 TeV	$\lvert\eta\rvert < 2.7$	$\lvert\eta\rvert < 2.4$

Another crucial part is the magnet systems. The ATLAS detector has two magnetic systems — the solenoid system and the toroid system. The solenoid system serves the inner detector with a 2 T magnetic field parallel to the beamline, while the toroid system provides 0.5 T and 1 T toroidal magnetic field with respect to the beamline for the muon spectrometers in central and end-cap region, respectively.

The ATLAS experiment uses right-handed coordinates. The nominal interaction point is defined as the origin. The x-axis points to the centre of the LHC ring; the y-axis points upwards;

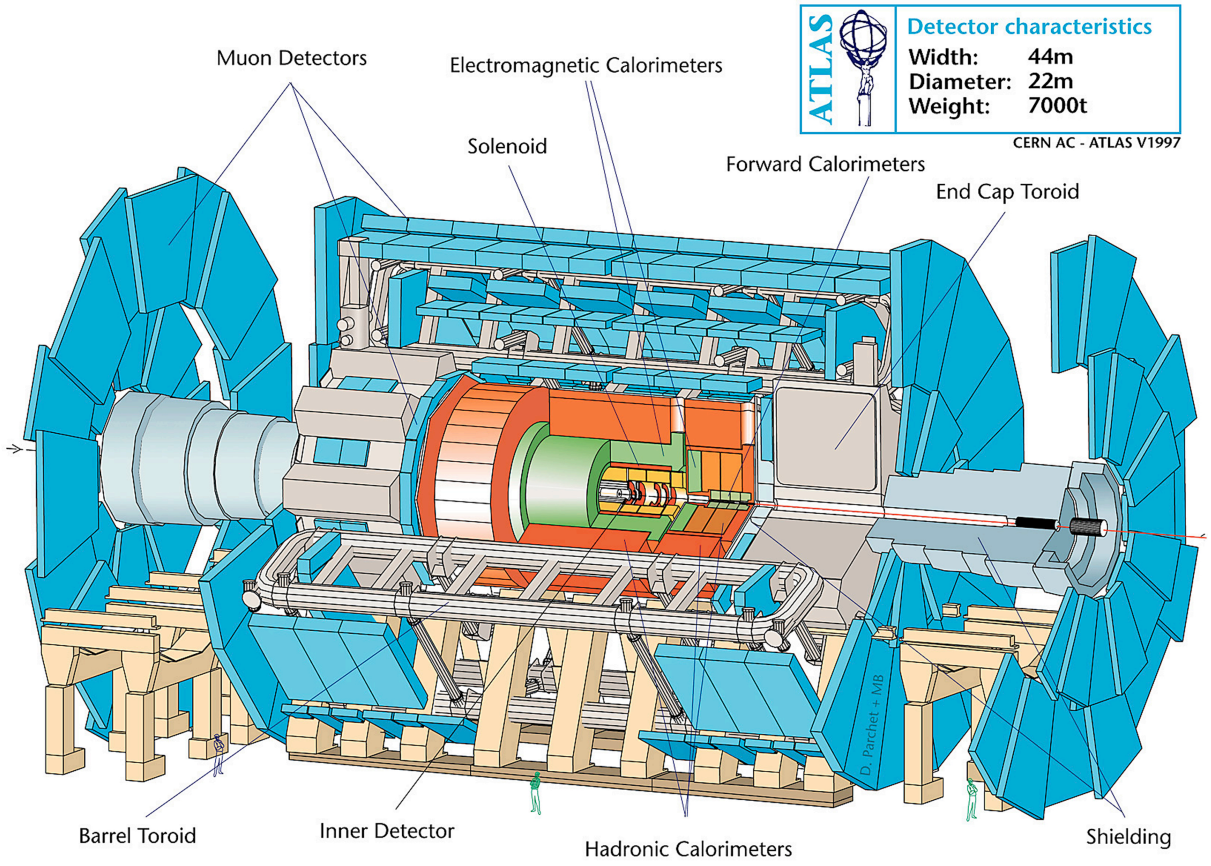


Figure 2.3: Demonstration of the ATLAS detector [57].

and the z-axis is aligned to the beamline. The azimuthal angle ϕ and the polar angle θ are defined based on this coordination system. More commonly, the pseudorapidity $\eta = -\ln(\tan\frac{\theta}{2})$ is used instead of θ to describe the Lorentz boost in the $z-$direction. Besides, a angular distance defined as $\Delta R = \sqrt{(\Delta\phi)^2 + (\Delta\eta)^2}$ is often used to measure the distance of two four-vectors in $\phi - \eta$ space.

2.2.1 Inner Detector

The Inner Detector locates closest to the beamline. Charged particles travel through the discrete layers of the inner detectors and interact with the material (so-called a hit), causing ionisation signal. The pattern of those hits is recognised to reconstruct the vertex and the track of the outgoing charged particles, which are used to determine the momentum, impact parameters and charge of these particles.

From the innermost to the outermost, the inner detector consists of three detector systems – the pixel detector, the silicon microstrip tracker (SCT) and the straw tube tracker called Transition Radiation Tracker (TRT).

The pixel detector is segmented in $R - \phi$ and z. Originally it has three layers in the barrel region. An additional Insertable B-Layer (IBL) [58] was installed before Run-2 to improve the primary and secondary vertex reconstruction. It also has three disks that are placed at both end-cap regions, perpendicular to the beamline. The SCT layouts similar as the pixel detector but extends

the radial measuring space. It consists of eight strip layers and nine disks. The TRT cannot render information on the z direction due to its nature (the straws are parallel to the z-axis). It detects transition radiations, which is particularly useful to identify electrons. The pixel detector and the SCT can detect tracks with a pseudorapidity of $|\eta| < 2.5$, while the TRT only covers tracks within $|\eta| = 2.0$.

2.2.2 Calorimeters

Next to the inner detector system are the sampling calorimeter detectors. Electrons, photons, and hadrons interact with the calorimeters and produce cascaded particles, called showers. The showers are categorised into electromagnetic ones and hadronic ones, depending on the original particle that initiates the shower. As the result of the showering, the energy of the particles is deposited in and absorbed by the calorimeter, therefore can be measured.

The electromagnetic calorimeter (ECal) is right next to the solenoid magnet that surrounds the inner detector. The active material of ECal is liquid-argon (LAr), with lead absorber and kapton electrodes plates designed in accordion-shape. It is segmented into a presampler layer and three ECal layers (EM1, EM2, EM3) with different designs. Table 2.2 summarises the typical thickness (in depth), typical granularity and main purpose of each layer [1]. The structure of ECal is sketched in Figure 2.4. In general, the thickness can make sure the energy of the shower is fully collected and the fine granularity can provide precise measurement of the energy of electrons and photons. The extra fine granularity of the first layer EM1 is used to discriminate neutral pion ($\pi^0 \to \gamma\gamma$) and photon. The segmentation in η is finer than the average separation of the photons from π^0 decay, which helps to distinguish the individual photons.

Table 2.2: The typical thickness, granularity and main purpose of each layers in the electromagnetic calorimeter system. The unit of thickness is mm or radiation length ($X_0 = 14.2$ cm) [46, 59].

Name	Thickness	Granularity ($\Delta\eta \times \Delta\phi$)	Purpose
Presampler	11 mm	0.025×0.1	Correct the energy loss ahead of the ECal.
EM1	$4.2X_0$	0.003×0.1	Discriminate neutral pion and photon.
EM2	$16X_0$	0.025×0.025	Collect the majority of the energy.
EM3	$2.0X_0$	0.05×0.025	Collect the energy of the tail of the shower.

The ECal covers the barrel and end-cap in the region of $0 < |\eta| < 1.475$ and $1.375 < |\eta| < 3.2$ (divided into two wheels that cover $1.375 < |\eta| < 2.5$ and $2.5 < |\eta| < 3.2$), respectively. The region $|\eta| < 2.5$ is dedicated to precision physics. The transition between barrel and end-cap is called the crack region, which is defined to be $1.37 < |\eta| < 1.52$.

Outside the electromagnetic calorimeter, there are also the hadronic calorimeters (HCal). It consists of the scintillator-tile calorimeter that is placed at the barrel region, the liquid-argon hadronic end-cap calorimeter (HEC) and the liquid-argon forward calorimeter (FCal). Plastic scintillators

[1] The thickness and granularity depends on η. Here we take the case of $\eta = 0$.

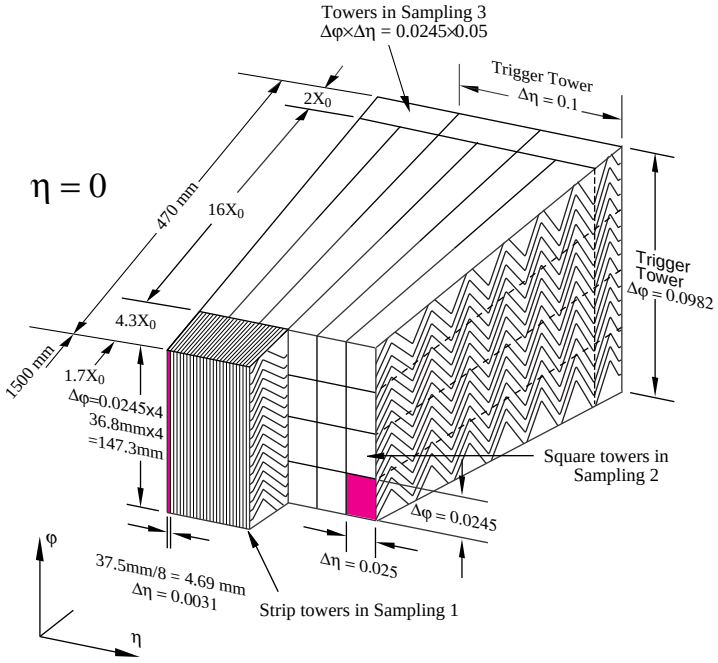


Figure 2.4: The accordion structure of the EM calorimeter [59].

are used as the active material and steel as the absorber in the tile calorimeter. The lights are collected by wavelength shifting fibres and are transported to photomultiplier tubes to measure the amount of energy of the hadronic shower. It has two parts that covers $|\eta| < 1.0$ and $0.8 < |\eta| < 1.7$. By its name, the HECs are installed in the end-cap region, covering $1.5 < |\eta| < 3.2$ It also utilises the LAr technique as ECal but uses copper as the absorber in a flat-plate design. The FCal covers the forward region to $|\eta| = 4.9$. It is designed to be able to measure both electromagnetic and hadronic showers, and to be dense to stop the radiation background in front of the muon spectrometer.

2.2.3 Muon spectrometer

The muon spectrometer (MS) lies in the outermost region of the ATLAS detector, where the momentum of muons whose trajectories are bended by the toroid magnets is precisely measured. The system contains two types of detectors. The Monitored Drift Tube (MDT) in the barrel region and the Cathode Strip Chamber (CSC) in the end-cap region cover $|\eta| < 2.7$, providing precise measurements on the tracks. The Resistive Plate Chamber (RPC) and the Thin Gap Chamber (TGC) that are installed in the barrel and the end-cap region serve as triggering chambers.

2.2.4 Trigger and data acquisition

The protons collide at a rate of approximately 1 GHz under a bunching crossing rate of 40 MHz. There is no technology to transfer and store all of the data. In fact, ATLAS reduces the rate of data recording to an average rate of 1 kHz using a two-level trigger system [60, 61]. A trigger and data acquisition (TDAQ) system is devoted to process, select and store the events for offline usage.

The Level-1 (L1) trigger is based on hardware. There are two types of L1 triggers. The L1 calorimeter trigger uses the information from the calorimeters. The information is processed by the Cluster Processor (CP) which selects the electron, photon and *tau*-lepton candidates, and the Jet/Energy-sum Processor (JEP) which selects the jet candidates and calculates the total transverse energy. The L1 muon triggers uses information from the muon trigger chambers to select muon candidates. The decision is made by the Central Trigger Processor (CTP) which combines the inputs from L1 calorimeter and L1 muon, as well as L1 topological processor (L1Topo) [62] and some other systems. The L1Topo is a processor with a number of algorithms that select events of interest by the kinematics and angular relation of the candidate objects. After this stage, the data rate is reduced to 100 kHz.

The data are then further processed and formatted to the High-Level Trigger (HLT), which is based on software. More detailed and precise algorithms run on a computer farm to make the trigger decisions. This stage is very similar to the offline physics analysis. To reduce the event rate, the HLT triggers usually require a high transverse momentum threshold of the objects or multi-objects. Finally, the data that passes the HLT requirements are stored for offline physics analysis.

2.3 Definition of physics objects

Collisions result in outgoing particles interacting with detectors, on the other hand, searches start with probing the final state particles. Defining the physics objects by the response of the detector is therefore a crucial step. This section gives an overview of the definition of tracks [63, 64], electrons [65, 66], muons [67, 68], jets [69, 70] and missing transverse energy ($E_{\mathrm{T}}^{\mathrm{miss}}$) [71] in the context of the search for $HH \to b\bar{b}\tau_{\mathrm{had}}^{+}\tau_{\mathrm{had}}^{-}$. Hadronic τ-leptons (τ_{had}) are not mentioned here because they are discussed in Chapter 3 dedicatedly.

2.3.1 Track and vertex

Tracks are reconstructed from the hits in the inner detectors fired by charge particles. They are the building blocks of the reconstruction, identification and calibration of the charged particles. Tracks are also used to reconstruct the primary and secondary vertex of a collision. It is particularly important for *b*-jet tagging.

The track reconstruction in the pixel detector and SCT consists of the following steps. [72]. Firstly, clusters of hits are created from pixels and strips. Next, the tracks are reconstructed by an iterative track finding based on the seeds built from the clusters. Multiple tracks per seed are built by the combinatorial approach [73]. Thus, in the last step, an ambiguity solver is utilised to compare and score the tracks, in order to choose the best track candidates .

The tracks are selected by the *Loose* quality criteria described in Ref. [64] except that the minimum p_{T} is increased to 500 MeV to handle the denser collision condition in Run-2. Primary

vertex is reconstructed by the vertex that has the largest sum of squared p_{T} over associated tracks. For physics analysis, events are required to have at least one primary vertex with at least two associated tracks.

2.3.2 Electron and Muon

Electrons leave trajectories in the inner detector and deposit energy in the ECal. The electron candidates are reconstructed by seed clusters of energy deposit that matches to the tracks fitted with the Gaussian-sum Fitter [74] in $\eta - \phi$ space. They are then identified by a likelihood-based (LH) identification (ID) requirement, which is constructed using variables of the tracks, the clusters and the matching between them.

Muons travel through the entire detector, deposit a small amount of energy in calorimeters and leave trajectories in the inner detector and the muon spectrometer. Based on different combinations of information from the inner detector, the calorimeters and the muon spectrometer, four types of reconstructed muons are defined [67], which are combined muons, segment-tagged muons, calorimeter-tagged muons and extrapolated muons. The muon candidates are identified by a set of requirements on the quality of the track reconstruction in each detector.

The electrons and muons that are produced in the processes of interest are usually isolated from the other particles in $\eta-\phi$ space while those from background processes are not. Applying isolation requirements can efficiently reject backgrounds. The isolation requirement can be defined using track-based or calorimeter-based information. In each case, the variable related to the isolation requirement is defined by the sum of momentum or energy of the tracks or clusters within a given cone around the electron or muon.

Selection requirements with different level of targeted signal efficiencies called working points are defined for the ID and isolation of electrons and muons. Details can be found in the corresponding references.

Table 2.3 summarises the selection requirements of electrons and muons in the search for Higgs boson pair production. In the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ decay channel, events with one or more such electrons or muons are vetoed.

Table 2.3: The requirements on the kinematics (p_{T}, η), identification and isolation of the electrons and muons in the search for $HH \to b\bar{b}\tau^{+}_{\mathrm{had}}\tau^{-}_{\mathrm{had}}$. The boosted analysis does not apply isolation criteria to electrons in order to reject boosted $Z \to ee$ background. For muons, the resolved analysis selects $|\eta| < 2.5$ to be harmonised with the search for $HH \to b\bar{b}\tau^{+}_{\mathrm{lep}}\tau^{-}_{\mathrm{had}}$ where muon triggers are utilised.

Variable	Electron	Muon
p_{T}	> 7 GeV	> 7 GeV
η	< 2.5 (not 1.37–1.52)	< 2.7 (see captain)
ID	*Loose*	*Loose*
Isolation	*Loose* (see captain)	*FixedCutLoose*

2.3.3 Jet

Jets are cones with a number of hadrons inside. The hadrons mainly consist of pions, and there also can be kaons, protons, neutrons and other hadrons. It is challenging to reconstruct the detector response of the hadrons or their decay products under the high pile-up condition. Jets are reconstructed using the anti-k_t algorithm [75] with radius parameter R. The reconstruction and calibration of jets are based on the inputs constructed by either the topological cell clusters (topo-clusters) [76] method or the particle flow (PFlow) [69] method.

The topo-cluster method uses the information provided by calorimeter cells. It generates topo-clusters by finding significant signal cells over pile-up induced noises. Initially, the hadronic clusters are calibrated at electromagnetic scale (EM scale) to give the same response as electromagnetic showers. To correct the hadronic clusters to the correct scale, the local cluster (LC) weighting scheme is employed.

The PFlow method combines the topo-clusters with the tracks in the inner detector. It can split the energy of different particles by subtracting the energy deposits of the charged particles in the calorimeters. After the subtraction, the particle flow objects (PFO) are defined. PFO includes the tracks of charged hadrons and the remaining part of the topo-clusters after the subtraction. Currently, the PFlow method still uses the EM scale for the topo-clusters. The PFlow method usually can provide more precise energy reconstruction.

Several types of jets are used in the search for Higgs boson pair production. In the resolved analysis, the jets from $H \to bb$ decay can be reconstructed individually. PFlow jets with $R = 0.4$ are used, namely small-radius(R) jets [69]. Jet cleaning algorithms are applied to remove jets from non-collision backgrounds and noises in the calorimeters [77]. The jets originate from pile-up are suppressed by the jet-vertex-tagger (JVT) [78].

Our analysis relies on the b-quark-initiated jets (b-jets). The b-jets are identified against the jets initiated by other sources by b-tagging algorithms. The b-tagging is performed mainly based on the distinctive secondary vertex created by the decay of long-lived b-hadrons. The algorithm is powered by a deep neural network called $DL1r$ [79, 80], which combines several low-level algorithms that extract information from associated tracks, primary and secondary vertices, and the reconstructed jet. The momentum of the b-tagged jets is further corrected by the μ-in-jet and PtReco methods described in Ref. [81]. These methods can improve the resolution of the invariant mass of two b-jets (m_{bb}). The resolutions of m_{bb} with and without the corrections are compared using the simulated $HH \to b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ non-resonant ggF+VBF HH samples (see Section 4.1) in Figure 2.5. The mass resolution is estimated by the full width at half maximum (FWHM) of the best-fit Bukin function [82]. It is found that the resolution improves by about 10%.

In the boosted analysis, two small-R jets are merged into one fat jet. In this case we use the LCTopo jets with $R = 1.0$, known as the large-radius(R) jets [70]. These jets are further trimmed to remove the effects of pile-up and underlying event. The trimming [83] is done by re-clustering the original constituents of a large-R jet into a collection of R_{sub} sub-jets using the k_t algorithm

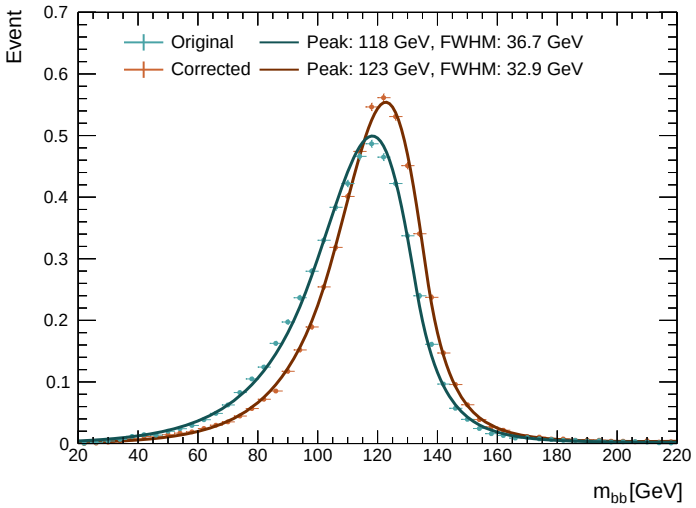


Figure 2.5: Comparison of m_{bb} resolution with and without the b-jet momentum corrections on the simulated $HH \to b\bar{b}\tau^{+}_{\text{had}}\tau^{-}_{\text{had}}$ non-resonant ggF+VBF signal samples.

[84–86]. The sub-jets are then discarded if their p_T is less than a specific fraction (f_{cut}) of the p_T of the original large-R jet. The values of the parameters for this analysis are $R_{sub} = 0.2$ and $f_{cut} = 5\%$. They are found to the optimal in Ref. [87]. The mass of the large-R jet is calculated by the combined mass technique [88]. The b-tagging of the large-R jet is performed with the variable-radius(R) track jets [70]. These jets are reconstructed from the inner detector using anti-k_t algorithm with a transverse momentum dependent radius parameter $R(p_{\text{T}}) = \rho/p_{\text{T}}$, the parameter ρ determines how fast the effective jet size decreases with the jet p_{T} [89]. Here ρ equals to 30 GeV and R is allowed to vary from 0.02 to 0.4 [90]. The variable-R track jets are assigned to large-R jets by the ghost association [91, 92]. The b-tagging of a large-R jet is determined by the b-tagging of the two leading variable-R track jets ordered by p_{T}. Here the b-tagging is based on a Boosted Decision Tree (BDT) method called $MV2c10$ [93], which is very similar to the $DL1r$.

Similar as the ID of the leptons, the b-tagging also provides working points with specific b-jet signal efficiencies. The selection requirements on the jets in the search for Higgs boson pair production are listed in Table Table 2.4.

Table 2.4: The requirements on the kinematics (p_{T}, η) and b-tagging of the jets in the search for HH. The b-tagging selection is defined by the working points based on b-tagging efficiency.

Variable	small-R jet	large-R jet	variable-R track jet
p_{T}	> 20 GeV	> 300 GeV	> 10 GeV
η	< 2.5	< 2.5	< 2.5
b-tagging	77% eff. ($DL1r$)	see variable-R track jet	70% eff. ($MV2c10$)

2.3.4 Overlap removal

Geometric overlap between reconstructed objects creates the chance of double counting. This situation is avoided by the following overlap removal requirements.

- Reject electron against muon if they share inner detector track. If the muon is a calorimeter-tagged muon, reject the muon.
- Reject small-R jet against electron if $\Delta R(\text{jet}, e) < 0.2$. For the remaining jets, reject the electron if $\Delta R(\text{jet}, e) < \min(0.4, 0.04 + 10\ \text{GeV}/p_{\text{T}}(e))$.
- Reject small-R jet against muon if those two conditions are satisfied (1) the muon Inner Detector track is ghost-associated to the jet or $\Delta R(\text{jet}, \mu) < 0.2$. (2) The jet has less than three associated tracks. For the remaining jets, reject the muon if $\Delta R(\text{jet}, \mu) < \min(0.4, 0.04 + 10\ \text{GeV}/p_{\text{T}}(\mu))$.

Additional overlap removal requirements are applied in the resolved analysis because of the τ_{had}'s:

- Reject τ_{had} against electron if $\Delta R(\tau_{\text{had}}, e) < 0.2$ and the electron passes *Loose* ID criterion.
- Reject τ_{had} against muon if $\Delta R(\tau_{\text{had}}, \mu) < 0.2$.
- Reject jet against τ_{had} if $\Delta R(\text{jet}, \tau_{\text{had}}) < 0.2$.
- Reject anti-τ_{had} (see Section 3.2.2) against jet if $\Delta R(\text{anti-}\tau_{\text{had}}, \text{jet}) < 0.2$ and the jet is b-tagged. For the remaining anti-τ_{had}'s, reject non-b-tagged jet is $\Delta R(\text{anti-}\tau_{\text{had}}, \text{jet}) < 0.2$.

2.3.5 Missing transverse energy

Missing transverse momentum $\vec{p}_{\text{T}}^{\,\text{miss}}$ is defined from the negative vector sum of the transverse momenta of all reconstructed visible objects [71]. In addition, soft terms from all tracks matched to the primary vertex but not associated with the higher-p_{T} reconstructed objects, are considered by the track-based soft term (TST) algorithm.

The $\vec{p}_{\text{T}}^{\,\text{miss}}$ can be written as

$$\vec{p}_{\text{T}}^{\,\text{miss}} = -\sum \vec{p}_{\text{T}}^{\,e} - \sum \vec{p}_{\text{T}}^{\,\mu} - \sum \vec{p}_{\text{T}}^{\,\gamma} - \sum \vec{p}_{\text{T}}^{\,\text{jet}} - \sum \vec{p}_{\text{T}}^{\,\text{soft}}, \tag{2.1}$$

The missing transverse energy ($E_{\text{T}}^{\text{miss}}$) represents the magnitude of $\vec{p}_{\text{T}}^{\,\text{miss}}$. In the search for HH in the $b\bar{b}\tau^+\tau^-$ decay channel, the main contribution to $E_{\text{T}}^{\text{miss}}$ is the neutrinos from the decays of τ-leptons.

2.4 Physics analysis with the ATLAS experiment

The physics analysis in this thesis is performed using the LHC Run-2 proton-proton collision data of an integrated luminosity of 139 fb^{-1} and a centre-of-mass energy of $\sqrt{s} = 13$ TeV. Only the data

identified as in good quality for physics analysis purpose are used. The list of datasets is provided by the Good Run List (GRL) [94].

To make predictions on the overall and differential composition of the data, simulated samples (a.k.a MC samples) are generated based on Monte Carlo technique. The generation of simulated events involves several steps [95]:

- Predict the distributions of the kinematics of the final state particles by calculating the Matrix Element (ME) of the hard interactions. Randomly generate events following the distribution. The tools for this step are called ME generators. The calculation is performed at a certain precision.
- Use phenomenological models provided by Parton Shower (PS) generators to simulate the parton shower and hadronisation of the final state particles come from the first step.
- Simulate how the outgoing particles from the last step interacts with the detectors and digitalise the response of the detector.

In the first step above, the underlying hard interaction in a p-p collision is the interaction of partons. Parton distribution functions (PDFs) are used to predict the energy fraction x_i carried by parton i. The evolution of PDF with energy scale is described by DGLAP equation [96]. PDFs are determined experimentally therefore uncertainties of the measurement needs to be considered. The α_s is also determined in experiments and it affects the measurement of PDFs, so the uncertainty of α_s is usually combined with that of the PDFs. A factorisation scale μ_F is defined to separate the scale to use PDF and the scale to calculate the ME of hard scattering. A renormalisation scale μ_R is set for the renormalisation of the strong coupling constant α_s. The nominal scale is set to the mass of Z boson m_Z. In ATLAS physics analysis, it is recommended to vary these two scales to estimate the uncertainty due to missing higher order corrections in the calculations of the DGLAP and the matrix element. The final interaction cross section depends on all the components above.

ATLAS minimum bias (MB) event and underlying event (UE) data are used to tune the parameter of the PS generators. The tunes to UE can be used in the simulation of UE of an interaction, and the tunes to MB is important for the simulation of pile-up. The pile-up is simulated separately with respect to the main interaction.

In ATLAS, the third step is done by GEANT4 [97] or Atlfast-II. The latter uses FASTCALOSIM package [98] to simulate the calorimeter response with parametrizations of the energy distributions, which reduces the simulation time by a order of magnitude, making it possible to quickly generate a specific sample. Samples simulated by GEANT4 are called Full-Simulation (FullSim) samples, while those simulated by Atlfast-II are called Fast-Simulation (FastSim) samples.

After these steps, the simulated samples are processed in the same way as the real collision data, passing through the physics object reconstruction routines to produce a data format for physics analysis. This format is called an Analysis-Oriented Dataset (AOD) format. To make precise predictions in downstream physics analysis, the simulated samples are calibrated to the data samples

at reconstructed object level (see Ref. [66, 68, 71, 99–101]). Besides, the amount of pile-up is reweighted to that of the real data, given the fact that the simulated samples are usually generated before the actual data-taking. Systematic uncertainties arising from the calibrations of the detector responses are propagated to the physics analyses. They are categorised as experimental uncertainties in the following.

The HLT and offline software of ATLAS experiment is contained in ATHENA [102]. Release version 22.x (R22) of ATHENA is used in the development of τ_{had} decay mode classification in Section 3.3, and the release version 21.x (R21) of ATHENA is used everywhere else.

Chapter 3

Performance of Hadronically Decaying τ-leptons

Reconstruction and identification of hadronically decaying τ-leptons (τ_{had}) are essential ingredients in the search for HH in the $b\bar{b}\tau^+\tau^-$ decay channel. This chapter outlines the current standard procedure of the reconstruction and identification and describes at length the design of the τ_{had} decay mode classification and its performance. As multivariate analysis (MVA) techniques are widely employed in this chapter and the next chapter, a brief introduction of the MVA is presented at the beginning. Finally, the reconstruction and identification of full-hadronic boosted di-τ objects are encapsulated.

3.1 Multivariate analysis

In the recent years, Machine Learning (ML) techniques have thrived rapidly and made a revolutionary impact to the society. Many tasks in the data analysis of particle physics can be converted into ML problems. Since decades, particle physics has adopted and developed ML-based algorithms to improve the operation of experiments and the analysis of data [103]. Alongside breakthroughs in ML, and in particular Deep Learning (DL), these applications in particle physics are superseded by even more advanced techniques, as reviewed by Ref. [104–106].

In particle physics, ML-based data analysis is also known as multivariate analysis (MVA). In a complicated data analysis (usually involving many variables), data analysis using the traditional method in which the algorithms are fully invented by our prior knowledges is sometimes impracticable. MVA method approaches the problem in a different direction, where the algorithms are learned by machine from data with much smaller needs for prior knowledges. Thus, it is a powerful tool to deal with complicated problems. Despite being criticised because the algorithms learned by MVA methods are hardly fully explainable, the promising improvements that they bring cannot be denied, as long as they are used with care.

In this thesis, MVA plays important role in the reconstruction of physics objects (this chapter) and the extraction of HH signal events (Section 4.2.4). To support later discussion, the basic

concepts of machine learning and two specific models used in this thesis are introduced below, referring to Ref. [107].

3.1.1 Basic concepts in multivariate analysis

Many problems of data analysis in particle physics, including those in this thesis, can be translated into supervised learning problems in ML. In these problems, an adaptive model $f_{\boldsymbol{w}}$ (vector $\boldsymbol{w}$ stands for the parameters of the model) is determined by a training set which includes both the input features (input variables) $\boldsymbol{x}$ and their paired ground truth (targets) $\boldsymbol{y}_{\text{truth}}$. There are two major types of supervised learning problems. Regarding the training set, if the vector $\boldsymbol{y}_{\text{truth}}$ is one or multiple discrete classes (categories), the problem is called a classification problem. If $\boldsymbol{y}_{\text{truth}}$ is one or multiple continuous values, the problem is called a regression problem.

The learning (training) phase is operated on the training set. In this phase, the parameters $\boldsymbol{w}$ of the model $f_{\boldsymbol{w}}$ are tuned by the $\{\boldsymbol{x}, \boldsymbol{y}_{\text{truth}}\}$ pairs. This process is repeated several iterations (epochs) until it is stopped by a specific condition, which is called an early stop. It is common to preprocess the input features to extract the most important information and make them less noisy for computing, such that the training can be faster and more stable.

After the training phase, the model is examined on an unseen dataset called the testing set, namely the inference phase. It may contain the ground truth $\boldsymbol{y}_{\text{truth}}$, for instance, if it consists of simulated events which contains all the generator level information. And it may not contain $\boldsymbol{y}_{\text{truth}}$, e.g., the testing set is the real collision data. In any cases, the goal is to find the optimal model which can make equivalently good predictions (labelled as $\boldsymbol{y}_{\text{pred}}$) of the $\boldsymbol{x}$ in both the training and the testing set. This is known as the generalisation of the model. If the performance with the training set is close to the ideal performance but the generalisation is poor, it usually indicates an overtraining issue. In our applications, the overtraining issue needs to be avoided as much as possible [105].

3.1.2 Methods

This subsection introduces two of the most popular machine learning methods used in particle physics experiments, the Boosted Decision Tree (BDT) [108, 109] and the Neural Network (NN) [110, 111].

3.1.2.1 Boosted decision tree

Decision trees are models that traverse a binary tree structure and perform selection on a chosen variable at a certain optimised threshold at each node along the path. The process recursively divides the space of the input feature vector $\boldsymbol{x}$ into cuboid regions, labelled as a vector $\boldsymbol{R}$. Once the partitioning stops, predictions $\boldsymbol{y}_{\text{pred}}^{R}$ are made for each region R. In classification problems, $\boldsymbol{y}_{\text{pred}}^{R}$ is a class, while in regression problems it is a constant value.

The learning of a decision tree is performed by constructing the tree structure in a greedy manner. Starting from the root node, the variable and the threshold for division are decided by finding the ones that provide the best performance at the current node. There are various ways to quantify the goodness of the performance., the metric is denoted as G. In a classification problem, the Gini index is often used. For a binary classification problem, it is

$$G = p \cdot (1 - p), \tag{3.1}$$

where p is the purity of one of the classes regarding the events in that node. In a regression problem, the average residual sum-of-square is used. It is,

$$G = \frac{1}{N} \sum_{i}^{N} (y^{i}_{\text{truth}} - \hat{y})^2, \tag{3.2}$$

where N is the number of events, and $\hat{y}$ is the average value of $\boldsymbol{y}_{\text{truth}}$ in that node. The growing of the tree is often stopped by requiring a minimum number of events in the node or maximum depth of the tree. The procedure can be easily extended to weighted datasets.

A single decision tree is well readable and explainable, but the performance is limited by the rigid partitioning approach. However, it can benefit from model combination methods, such as *boosting*. For example, the signal-background discrimination in Section 4.2.4 uses the *gradient boosting* algorithm. This algorithm defines a differentiable loss function $L(y, F(\boldsymbol{x}))$ (note that the subscript 'truth' of y is omitted) [1], where $F(\boldsymbol{x})$ is the weighted sum of M decision trees, which can be written as

$$F(\boldsymbol{x}) = \sum_{m=1}^{M} \gamma_m h_m(\boldsymbol{x}) + const. \tag{3.3}$$

Here, the γ_m are the weights and the h_m are the decision trees. The loss function can be interpreted as a negative log-likelihood function. The goal is to find the $F(\boldsymbol{x})$ that minimises the loss function. This algorithm starts at a constant value, and then iteratively

- Calculate the pseudo-residual for each event, which is

$$r_{im} = -\left[\frac{\partial L(y_i, F(\boldsymbol{x}_i))}{\partial F(\boldsymbol{x}_i)}\right]_{F=F_{m-1}}, \tag{3.4}$$

  where i is the index of events (this is the 'gradient' of gradient boosting);
- Use a decision tree $h_m(\boldsymbol{x})$ to estimate the pseudo-residual;
- Find the γ_m that minimises the loss function at that iteration;
- Update $F(\boldsymbol{x})$: $F_m(\boldsymbol{x}) = F_{m-1}(\boldsymbol{x}) + \gamma_m h_m(\boldsymbol{x})$.

After M iterations, the final model $F(\boldsymbol{x})$ is learned, which is called a BDT model. The BDTs used

[1] In TMVA implementation it is $\ln\left(1 + e^{-2F(\boldsymbol{x})y}\right)$.

in this thesis are implemented by the TMVA [112] package. For the BDT in Section 4.2.4, the input y is -1 for backgrounds and $+1$ for signals, therefore the output of BDT, also called a BDT score or a BDT response, is between -1 (stands for background-like) and $+1$ (stands for signal-like).

3.1.2.2 Neural network and deep learning

Neural networks (NNs) are models that imitate the mechanism of human's neurons and synapses based on simple linear models together with non-linear operations. In the first layer (known as the input layer) of the NN, the input D-dimensional vector $\boldsymbol{x}$ is linear transformed into

$$a_j = \sum_{i=1}^{D} w_{ji}^{(1)} x_i + w_{j0}^{(1)}, \text{ where } j = 1, \ldots, M. \tag{3.5}$$

The transformed vector $\boldsymbol{a}$ is a M-dimensional vector. The parameters $w_{ji}^{(1)}$ and $w_{j0}^{(1)}$ are called weights and biases, respectively. The superscript (1) indicates that they are the parameters of the linear transformation acting on the first layer (input layer) of the NN. A differentiable, non-linear function $h(\cdot)$ called activation function acts on each a_j and produces

$$z_j^{(1)} = h(a_j), \tag{3.6}$$

which are called hidden nodes. Multiple hidden nodes form a hidden layer of the NN, supposing that there are M hidden nodes. Using $z_j^{(1)}$ as input, another hidden layer can be derived

$$z_k^{(2)} = h\Big(\sum_{i=1}^{M} w_{kj}^{(2)} z_j^{(1)} + w_{j0}^{(2)} \Big). \tag{3.7}$$

The process can be repeated until the final output layer with $\boldsymbol{y}$ is defined. Neural networks constructed in this way are called feed forward neural networks. With the development deep learning technique, the neural networks may have many hidden layers (deep neural networks). Various architectures are invented apart from the most common feed forward NNs. For instance, the recurrent neural networks used in b-jet tagging and τ_{had}-ID are NNs that operate on a sequence of input vectors. The sequence is looped over in the forward information propagation so that the hidden layer is calculated based on all the previous steps. The weights and biases are shared for each step of the sequence.

For the hidden layers, a popular choice of activation function is *ReLU* function [2]. For the output layer, no activation function is needed for regression problems, while *logistic (sigmoid)* function [3]

[2] The ReLU activation function is $\text{ReLU}(a) = \begin{cases} a, & a \geq 0 \\ 0, & a < 0 \end{cases}$

[3] The logistic activation function is $\sigma(a) = \frac{1}{1+e^{-a}}$, which can be considered as a special case of softmax function.

and *softmax* function [4] are used for two-class (binary) and multi-class classification problems. For the classification problems, the outputs are considered as the probabilities of being in that class. Loss functions (Error functions) $E(\boldsymbol{w})$ are negative log-likelihood functions of the weights and biases $\boldsymbol{w}$ defined based on these probabilities. For regression problems, it is

$$E(\boldsymbol{w}) = \frac{1}{2}\sum_{n=1}^{N} \left\| \boldsymbol{y}_n(\boldsymbol{w}) - \boldsymbol{y}_{\text{truth}} \right\|. \tag{3.8}$$

The summation over N is usually operated on part of the dataset called a batch, the value of the loss is thus calculated per batch. For binary classification problems, it is

$$E(\boldsymbol{w}) = -\sum_{n=1}^{N} \left[y_{n,\text{truth}} \ln y_n(\boldsymbol{w}) + (1 - y_{n,\text{truth}}) \ln(1 - y_n(\boldsymbol{w})) \right], \tag{3.9}$$

where $y_{n,\text{truth}}$ is 0 or 1. In the multi-class classification problems, the loss function is

$$E(\boldsymbol{w}) = -\sum_{n=1}^{N}\sum_{c=1}^{C} y_{cn,\text{truth}} \ln y_{cn}(\boldsymbol{w}), \tag{3.10}$$

where c sums over the C categories. The vector $\boldsymbol{y}_{n,\text{truth}}$ ($y_{cn,\text{truth}}$ are the elements of the vector) is a C-dimensional vector called a one-hot label, i.e. If the truth category is c, then only $y_{cn,\text{truth}} = 1$ and the values of the rest of the elements are all zero.

The goal of the NN algorithm is to find the parameters $\boldsymbol{w}$ that minimise the loss function. Optimisation algorithms are developed to meet this purpose, which try to make small steps in the opposite direction of the gradient of the loss function ($-\nabla E(\boldsymbol{w})$) per batch to get closer to the global minimum value as much as possible. The gradients of a feed forward NN are evaluated by the error back-propagation technique which is based on the chain rule of derivatives. In practise, the optimisation problem can be highly complicated. Therefore, finding a small enough local minimum is considered to be sufficiently good. The parameters $\boldsymbol{w}$ are updated in the optimisation process. In the most common *stochastic gradient descent (SGD)* algorithm, the weights and biases are updated in the form

$$\boldsymbol{w}^{\tau+1} = \boldsymbol{w}^{\tau} - \eta \cdot \nabla E(\boldsymbol{w}^{\tau}), \tag{3.11}$$

based on a randomly sampled mini-batch from the current batch, where τ indicated the index of current step. The parameter η is called learning rate, which scales the size of the update. More advanced optimisation algorithm finds the minimum faster, e.g. *Adam* algorithm, which is based on SGD but also considers the updates in the previous iteration by a *moment*, and the size of the moment is self adaptive.

The value of the loss function is monitored during the training, presented as a function of

[4] The softmax activation function is $\sigma(a_i) = \frac{e^{a_i}}{\sum_{j=1}^{C} e^{a_j}}$ for the i-th category, where C is the number of categories.

epoch [5]. After each epoch of training, the value of loss is calculated on an independent dataset called validation set. The value is expected to first decrease and then increase when the over-training starts. For the purpose of generalisation, the training is stopped when the validation loss reaches its minimum.

3.2 Hadronic τ object

3.2.1 Properties of τ-leptons

Tau-lepton (τ-lepton) is the third-generation lepton in the standard model, with a mass of 1.777 GeV and a proper lifetime of 290 fs, i.e., proper decay length of 87 μm [23]. It either decays leptonically ($\tau \to \ell\nu_\ell\nu_\tau, \ell = e, \mu$, labelled as τ_{lep}) or hadronically ($\tau \to$ hadrons ν_τ, labelled as τ_{had}). The τ_{lep} and τ_{had} represent the visible components of the decay products. In ATLAS detector, the decay typically happens inside the beam pipe. Only hadronic τ decays are considered in this chapter. It contributes about 65% to the total τ decays. The decay products of τ_{had} are dominantly consists of one or three charged particles. They present about 77% and 23% of all τ_{had} decays, respectively. In the following, they are labelled as *1-prong* (or *1P*) and *3-prong* (or *3P*) τ_{had}'s, indicating the number of charged particles in the decay products. Five major decay modes of τ_{had} and their branching ratio with respect to all of the hadronic τ decays are summarised in Table 3.1. The labels for each decay modes are shown for later discussion in Section 3.3. They are in form of 'ApBn', where 'A','B' are the number of prongs (the 'p') and neutral pions (the 'n') and 'X' stands for more than one or more than zero accordingly.

Table 3.1: Five major decay modes of τ_{had} and the corresponding branching ratio (BR) with respect to total τ decays [23]. $h^\pm$ stands for $\pi^\pm$ or $K^\pm$ (the latter contributes about 3%). Neutrinos are omitted. Decays with neutral kaons are excluded.

Decay mode	Label	BR (w.r.t total τ decays)
$h^\pm$	1p0n	11.5%
$h^\pm\pi^0$	1p1n	25.9%
$h^\pm \geq 2\pi^0$	1pXn	10.6%
$3h^\pm$	3p0n	9.46%
$3h^\pm \geq 1\pi^0$	3pXn	5.09%

3.2.2 Reconstruction and identification

The reconstruction of τ_{had} defines the coordinate system for τ_{had} variables, the number of prongs, electric charge and four-momentum [6] of τ_{had}. In the identification stage, τ_{had} is discriminated from

[5] Here, an epoch is the number of iterations over the entire training set

[6] The mass of τ_{had} is defined as 0.

backgrounds from quark-/gluon-initiated jets and electrons. The full reconstruction and identification chain of τ_{had} is briefly introduced below.

The τ_{had}'s in ATLAS experiments are seeded from the anti-k_t ($R = 0.4$) topo-cluster jets with LC energy correction [113]. A minimum p_{T} of 10 GeV and $|\eta| < 2.5$ is required for the seeding jets. On this stage, the energy of τ_{had} is the total energy of topo-clusters within the $\Delta R < 0.2$ cone with respect to the seeding jet direction. It is called the LC scale energy. The topo-clusters consist of two components in τ_{had} reconstruction. The hadronic component contains the calorimeter cells in the last layer of the electromagnetic and the hadronic calorimeters; The electromagnetic component contains those in the other layers of the ECal.

To improve the reconstruction in the high pile-up environment, the τ vertex is defined to be the primary vertex candidate that has the largest fraction of the summed momentum of track candidates [7] in the $\Delta R < 0.2$ cone around the seeding jet direction [113]. The variables for τ_{had} and τ_{had} tracks are derived based on the coordinate system built by this τ vertex. For instance, the direction (η or ϕ) of τ_{had} is defined by the vector sum of the topo-clusters within the $\Delta R < 0.2$ region using the τ vertex as origin. The reconstruction of τ_{had} using τ vertex is more efficient than using the primary vertex at low p_{T} and high pile-up [113].

The track candidates with the $\Delta R < 0.4$ region around the direction of τ_{had} is classified into four categories using a BDT-based algorithm based on the variables related to the transverse momentums, directions, impact parameters, number of hits in different layers of the inner detector and probability of being a track from electron [114]. The four categories are

- τ_{had} *tracks*: from the charged particle(s) in τ decay;
- *isolation tracks*: from underlying events of the p-p collision;
- *conversion tracks*: from photon conversion process ($\gamma \to e^+e^-$);
- *fake tracks*: from other sources, e.g., from pile-up.

The reconstructed (classified) τ_{had} tracks are used to define the electric charge and number of prongs of the τ_{had} that the tracks are associated with. The others types are not directly used in the physics analysis, but they are useful for the τ_{had} identification and τ_{had} decay mode classification that are discussed later. For 1-prong τ_{had}'s, the rate of correct reconstructing the number of prongs is about 85%, while for 3-prong τ_{had}'s it is 30%–70% depending on the p_{T} of τ_{had}. The performance for 3-prong τ_{had} degrades at high p_{T} because the tracks are so collimated that it is difficult for the detector to resolve individual tracks.

There are two calibrations of τ_{had} energy available, the baseline calibration and the Boosted Regression Tree (BRT) calibration [101]. Their goal is to correct the measured energy in the detector to the real generator-level energy. In the baseline calibration, the contribution from pile-up is subtracted from the energy at LC scale, and the corrected energy is divided by a response factor calculated from the Gaussian mean of the ratio of that and the generator-level visible decay prod-

[7] Tracks with $p_{\text{T}} > 1$ GeV and satisfy quality criteria based on the number of hits in the inner detector (≥ 2 hits in pixel detector and ≥ 7 hits in pixel and SCT detectors).

ucts of τ-lepton. The energy resolution of the baseline calibration is good for high p_{T} τ_{had} but poor for low p_{T} τ_{had}, because it only use the energy deposits in the calorimeters. The low p_{T} performance is largely improved by the τ_{had} particle flow reconstruction method [115]. In this method, the individual charged and neutral hadrons in the τ decay can be reconstructed, and the momentum of charged hadrons are from the significantly more precise measurement in the inner detector. The τ_{had} particle flow method will be discussed further in Section 3.3. A BRT is trained based on the information from both the baseline calibration and the particle flow calibration plus a few more variables from the topo-clusters and tracks to predict the true energy of τ_{had}. The final prediction is called BRT-based energy scale. It is the energy scale used by the physics analyses in this thesis. Thanks to the information from the particle flow calibration, the energy resolution of the BRT calibration is about twice better than the baseline calibration at lower p_{T} ($p_{\mathrm{T}} < 100$ GeV) region.

The main background of τ_{had} is jets originates from quarks and gluons, the above reconstruction only provides limited background rejection ability. Dedicated identification algorithm is developed to discriminate τ_{had} and these jets, namely τ_{had} identification, or τ_{had}-ID. In Ref. [113], an BDT-based τ_{had}-ID is presented, which is based on several variables constructed with the tracks and the electromagnetic and hadronic components of the topo-clusters. These variable describe the energy factions, electromagnetic shower shape, the π_0 component in the τ decay, etc. During Run-2, a Recurrent Neural Network (RNN) τ_{had}-ID is developed, where the features of each track and topo-cluster that are associated to the τ_{had} are considered in the MVA model are added for the classification [116]. The RNN learns additional information from sequences from tracks and sequences of topo-clusters, therefore it provides better performance compared with the BDT τ_{had}-ID. The improvement in background rejection is about 75%–100% depending the τ_{had} p_{T}. In both methods, the models are trained and evaluated on 1- and 3-prong τ_{had}'s separately. Four working points (*VeryLoose*, *Loose*, *Medium*, *Tight*) are provides for both 1-prong and 3-prong RNN τ_{had}-IDs, with baseline efficiencies of (95%, 85%, 75%, 60%) and (95%, 75%, 60%, 45%) in that order, respectively. At the *Loose* working point, the jet background mis-identification efficiencies are 0.01 and 0.05 for 1- and 3-prong τ_{had} candidates. The identification efficiency is stable against τ_{had} p_{T}, η and pile-up.

Another BDT-based model is employed to discriminate 1-prong τ_{had}'s from electrons, which is called a eVeto [117]. It uses the information from the TRT, the angular distance between track and τ_{had} direction and energy fractions in different components of the calorimeter. Three working points are provided. They are *Loose*, *Medium* and *Tight*, which correspond to signal efficiency of 95%, 85% and 75%, respectively. Contamination from muons can be reduced by the muon identification [117]. It is sufficient for the physics analyses in this thesis.

Basic requirements that are applied on τ_{had} in the search for HH are summarised in Table 3.2. The τ_{had} that fails the *Loose* ID working point and passes the *VeryLoose* ID working point is called a anti-τ_{had}. They are used in the estimation of backgrounds that contain misidentified τ_{had}'s.

Table 3.2: The requirements on the kinematics (p_T, η) and identification of τ_{had}'s in the search for HH. Selection on τ_{had}-ID is inverted to define anti-τ_{had}. 'n-prong' stands for the number of prongs (τ_{had} tracks associated to τ_{had}).

Variable	τ_{had}	anti-τ_{had}
p_T	> 20 GeV	
n-prong	1 or 3	
$\lvert\eta\rvert$	< 2.5 (not 1.37–1.52)	
τ_{had}-ID	*Loose*	*VeryLoose* not *Loose*
eVeto	*Loose*	

3.3 Decay mode classification of τ_{had}

3.3.1 Introduction

The standard reconstruction procedure of τ_{had}'s described in the last subsection only reconstructs the number of charged hadrons in the τ_{had} decay. However, various physics analyses require information of the neutral hadrons as well. For example, the measurement of CP mixture of the Yukawa coupling between Higgs boson and τ-lepton via the spin effects in the $H \to \tau\tau$ decay relies on precise identification of τ_{had} decay modes [118, 119], especially $h^{\pm}$, $h^{\pm}1\pi^0$ and $3h^{\pm}$ (see Table 3.1).

ATLAS has developed a method called the τ_{had} Particle Flow (TPF) reconstruction [115], which can be used to reconstruct the individual charged and neutral hadrons in the τ_{had} decays. This method provides a way to classify the five primary decay modes in Table 3.1 and reconstruct the momentum of τ_{had} using these reconstructed hadrons. It significantly improves the energy resolution of τ_{had} with p_T lower than around 100 GeV [8].

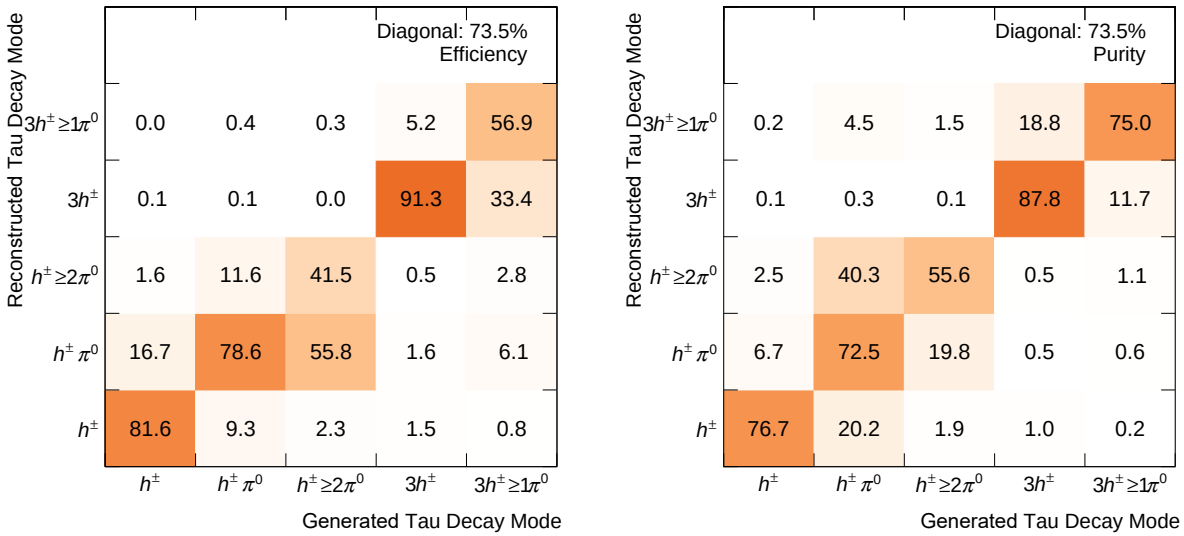
In the TPF reconstruction, the τ_{had}'s with p_T between 15 GeV and 100 GeV is mainly considered. That is the energy scale of τ_{had} p_T in most of the electroweak and SM Higgs boson. The typical p_T of the hadrons is about 4 GeV and with an average separation of $\Delta R \simeq 0.07$ [115].

The charged hadrons are reconstructed from the tracks in the inner detector, which provides significantly better energy resolution than the calorimeters at lower p_T scale, as indicated in Table 2.1. The origin of the tracks is classified by the method described in Section 3.2.2. Charged hadrons ($h^{\pm}$) correspond to the reconstructed τ_{had} tracks. The main misclassification and contamination are due to the conversion tracks produced by the $\gamma \to e^+e^-$ background. The $h^{\pm}$'s deposit a small fraction of energy in the ECal and most energy in the HCal, with highly irregular shower shapes.

The neutral hadrons (π^0's) which immediately decay to photon pairs. Consequently, the π^0's are reconstructed by the energy deposits in the ECal created by the photons. The separation between hadrons is similar to the width of the π^0 shower in the ECal, which makes it difficult to disentangle energy deposits from π^0 and $h^{\pm}$. The π^0 candidates are created by clustering ECal cells around

[8] The energy scale is used as input to the BRT τ_{had} energy scale algorithm (see Section 3.2.2).

the τ_{had}. The energy of π^0 candidates is corrected by subtracting the energy deposits of the nearest reconstructed $h^{\pm}$'s in the ECal. Many π^0 candidates come from $h^{\pm}$ remnants, pile-up process and other sources. Minimum p_{T} selection (around 2.1–2.7 GeV) and identification criterion are applied to improve the purity of π^0. The identification is based on a BDT, using information from the topo-clusters (e.g. energy density, shower shape)[9]. The counting of $h^{\pm}$ and π^0 provides a preliminary decay mode classification. However, about half τ_{had}'s in $h^{\pm} \geq 2\pi^0$ mode are misidentified as $h^{\pm}\pi^0$. In these cases, the photons from the π^0's can not be resolved, such that only one cluster is built. The information from EM1 layer of ECal (see Section 2.2.2) is used to improve the $h^{\pm} \geq 2\pi^0$ reconstruction. Local energy maxima of the energy deposits in EM1 around the τ_{had} are searched. The number of local maxima associated with a π^0 candidate is counted based on the transverse energy E_{T} of it [10]. Finally, the kinematic and π^0 identification information from the selected $h^{\pm}$ tracks, π^0 candidates and local maxima in EM1 are combined using a BDT-based decay mode classification algorithm called PanTau. PanTau utilises three BDTs to discriminate $h^{\pm}$ against $h^{\pm}\pi^0$, $h^{\pm}\pi^0$ against $h^{\pm} \geq 2\pi^0$ and $3h^{\pm}$ against $3h^{\pm} \geq 1\pi^0$. Figure 3.1 illustrates the efficiency matrix that presents the probability for each generated decay mode (x-axis) to be reconstructed as a particular decay mode (y-axis), and the purity matrix that shows the fraction of generated true decay mode (x-axis) in each reconstructed decay mode (y-axis). The performance is evaluated using the same simulated $\gamma^* \to \tau\tau$ samples and selections introduced in Section 3.3.2.



(a) Efficiency (b) Purity

Figure 3.1: Efficiency (a) and purity (b) matrices of PanTau evaluated on simulated $\gamma^* \to \tau\tau$ samples.

The TPF reconstruction was developed using Run-1 simulated samples and calibrated using Run-1 5 fb^{-1} data collected at $\sqrt{s} = 8$ TeV [115]. The algorithm is re-optimised using Run-2 sim-

[9] Many of the input features are also used in the decay mode classification (see Section 3.3.2.1).

[10] Counted twice if $E_{\text{T}} > 10$ *GeV*) [115].

ulated samples to gain the best performance for the studies in the current data-taking period. The four-momentum reconstruction using TPF approach can be found in Ref. [115], where the $h^{\pm}$ and π^0 candidates summed according to the reconstructed decay mode. As mentioned in Section 3.2.2, the TPF energy scale is used by the BRT energy scale.

3.3.2 Improvements in decay mode classification

The BDT-based PanTau decay mode classification outperforms the counting-based approach by about 3.8% in terms of diagonal efficiency/purity. However, it does not exploit the information from all the relevant objects that are associated with the τ_{had}. To fully capture the complexity and further improve the performance of the classification, a particular type of Graph Neural Network (GNN) [120] called the DeepSet neural network (DeepSet NN) [121] is utilised. The DeepSet is able to consider all the objects associated with the τ_{had} and process their information in a permutation invariant manner. The architecture and training of DeepSet NN for decay mode classification are detailed in Section 3.3.2.2. The input features that are considered for each objects are presented in Section 3.3.2.1. Finally, in Section 3.3.3, the expected performance of DeepSet NN classification is summarised.

The improved model supersedes the PanTau BDTs in the future Run-3 reconstruction software, e.g., starting from version R22 of ATHENA. The underlying reconstruction algorithms (track and jet reconstruction, etc.) that the decay mode classification is based on are not expected to make a significant impact, thus the changes in these underlying algorithms are not covered here. Dedicated samples are produced by R22 reconstruction software for the training of the DeepSet NN and evaluation of both the PanTau and the DeepSet NN. The $\gamma^* \to \tau\tau$ events are generated using PYTHIA 8 as the ME and PS generator with A14 [122] set of parameters tunes and NNPDF2.3LO [123] as the PDF. It also handles the generation of the underlying events and τ-lepton decays. Only fully hadronically decaying $\tau\tau$ are considered in the event generation. Unlike the physical Drell-Yan process $Z/\gamma^* \to \tau\tau$, the $\gamma^* \to \tau\tau$ process is generated with a smooth and flat invariant mass spectrum of m_{γ^*} and consequently, provides a smooth and flat spectrum of τ_{had} p_{T}. The beam spot weights [124] are considered in both training and evaluation to account for the difference of the beam spot width in data and simulated samples. The PanTau BDTs are not re-optimised using the R22 samples, but it is found that the performance of PanTau is consistent between the simulated samples produced using R21 and R22 reconstruction facilities, because it is stable against pile-up by nature.

3.3.2.1 Input features for classification

According to the TPF reconstruction, four sets of objects that are associated with τ_{had}'s are important in the decay mode classification:

- The $h^{\pm}$ that are the tracks in τ_{had} decay, called reconstructed τ_{had} tracks;

- The selected π^0 candidates that can come from the τ_{had} decay or $h^{\pm}$ remnants, pile-up, etc., called neutral particle flow objects (neutral PFOs);
- The local energy maxima in the EM1 layer of ECal associated with the π^0 candidate, called photon shot particle flow objects (shot PFOs);
- The reconstructed conversion tracks produced by $\gamma \to e^+e^-$.

The kinematic variables of the four types of objects are used as input features. Here, 'object' refers to a τ_{had} track, neutral PFO, shot PFO or conversion track. The p_{T} of the τ_{had} [11] and the object are used, which is transformed to their logarithm values ($\log_{10}$). The angular distances between the object and the τ_{had} are also considered. Besides, the angular distance between the object and the extrapolated direction of the leading-p_{T} τ_{had} track into the ECal is considered as an extra information. It accounts for the fact that the direction of τ_{had} fully rely on the calorimeter information, but the τ_{had} track (hadrons) and conversion tracks (electrons) undergo different interactions with the ECal. The name of the variables are summarised in Table 3.3. The variables can be visualised by a τ_{had} decay display as shown in Figure 3.2 (only $\Delta\eta(\cdot, \tau_{\text{had}})$ and $\Delta\phi(\cdot, \tau_{\text{had}})$, for $\Delta\phi(\text{object}, \text{trackECal})$ and $\Delta\eta(\text{object}, \text{trackECal})$ see Figure A.1). The prime in $\Delta\eta'$ and $\Delta\phi'$ indicates that they are transformed from the original $\Delta\eta$ and $\Delta\phi$ for better visualisation, where $\Delta x' = \sqrt{0.4^2 - (0.4^2 - x^2)}$, $x = \eta, \phi$. The circles represents the relative position of the object in the $\Delta\eta'$–$\Delta\phi'$ plane, with radius proportional to $\sqrt{\frac{p_{\text{T}}(\text{object})}{p_{\text{T}}(\tau_{\text{had}})}}$. Difference types of objects are displayed in different colours. As reference, the grey rings shows the original $\Delta R = \sqrt{\Delta\eta^2 + \Delta\phi^2}$ at various values. We can assume that the DeepSet NN is able to learn the topology of the objects as what is displayed. In fact, it is found that the preliminary DeepSet NN which only takes the kinematic variables already outperforms the PanTau model by about 3% in terms of diagonal efficiency/purity.

Table 3.3: Name and description of the kinematic variables as input features to the DeepSet NN decay mode classification. 'Object' refers to a τ_{had} track, Neutral PFO, Shot PFO or conversion track. The τ_{had} is the one that the objects are associated to by the TPF reconstruction. 'TrackECal' refers to the extrapolation of leading-p_{T} τ_{had} track to ECal.

Name	Description
$\Delta\phi(\text{object}, \tau_{\text{had}})$	Distance between the object and τ_{had} in ϕ
$\Delta\eta(\text{object}, \tau_{\text{had}})$	Distance between the object and τ_{had} in η
$\Delta\phi(\text{object}, \text{trackECal})$	Distance between the object and the extrapolation of leading-p_{T} τ_{had} track to ECal in ϕ
$\Delta\eta(\text{object}, \text{trackECal})$	Distance between the object and the extrapolation of leading-p_{T} τ_{had} track to ECal in η
$p_{\text{T}}(\tau_{\text{had}})$	p_{T} of τ_{had}.
$p_{\text{T}}(\text{object})$	p_{T} of the object.

For the (τ_{had} and conversion) tracks, the impact parameters d_0 and $z_0 \sin\theta$ are utilised. The corresponding impact parameter significance $Sig(d_0)$ and $Sig(z_0 \sin\theta)$ are also exploited, which are the

[11] Here the p_{T} of τ_{had} are defined by the baseline energy scale.

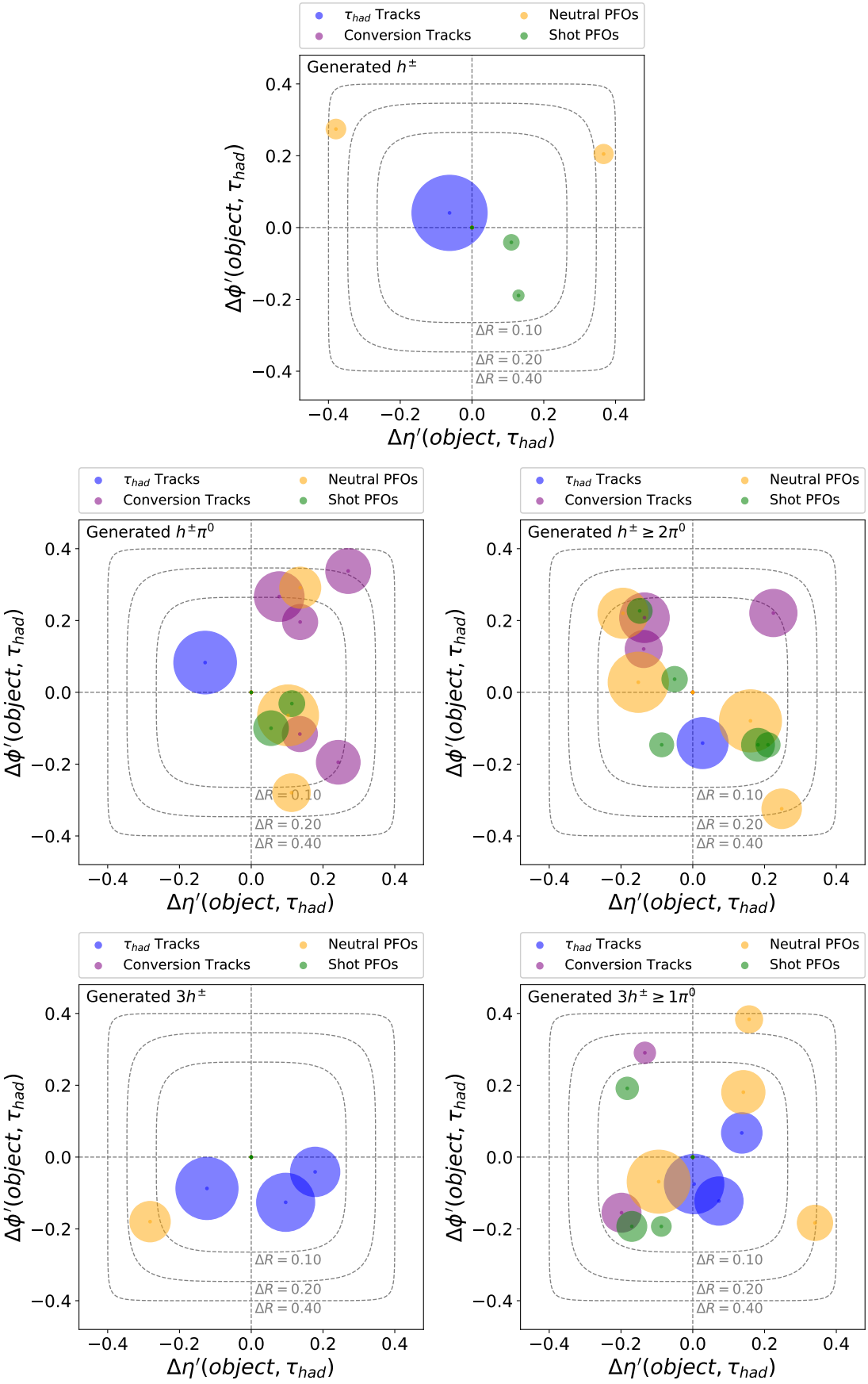


Figure 3.2: Visualisation of kinematic variables in displays of various generated τ decay modes with the simulated $\gamma^* \rightarrow \tau\tau$ sample. Details on the figures are explained in the text.

ratios of the impact parameters to their resolutions. The parameter d_0 is the closest distance of the track to the τ_{had} vertex in the transverse plane and $z_0 \sin\theta$ is the longitudinal closest distance. The distribution of the impact parameters and impact parameter significances of the leading-p_{T} tracks in different decay modes is shown in Figure 3.3. These variables provide about 2% improvements in the diagonal efficiency/purity.

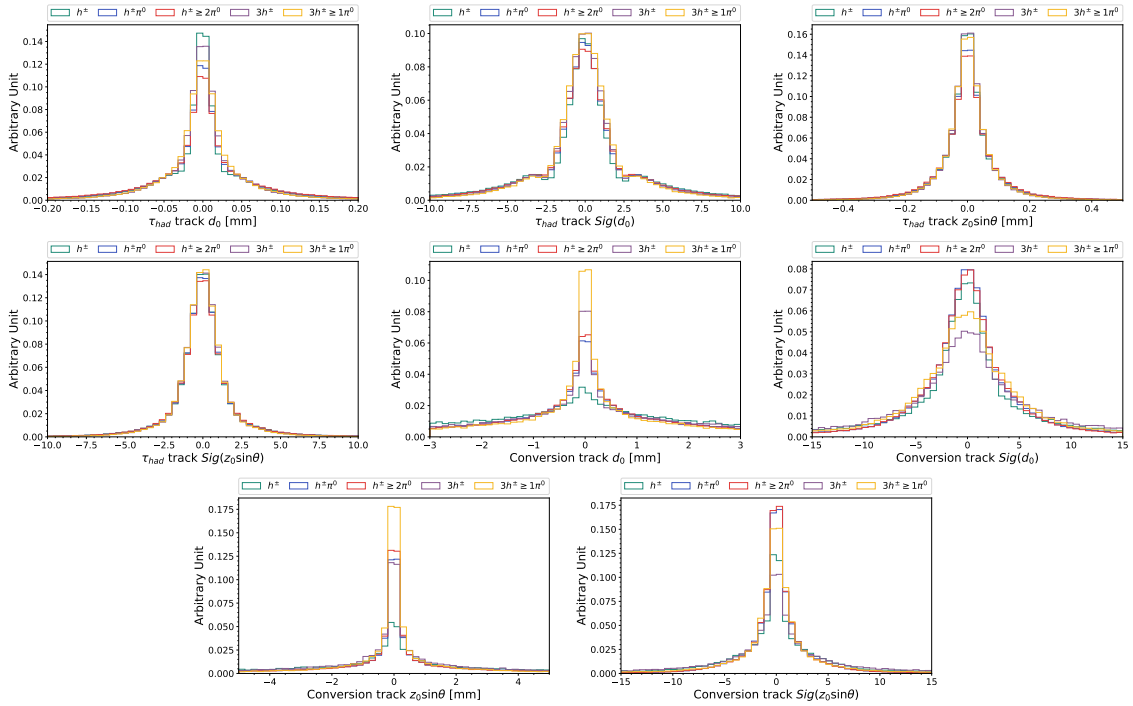


Figure 3.3: Distribution of the impact parameters and impact parameter significances of the leading-p_{T} tracks in different decay modes, normalised to unit integral.

For the neutral PFOs, the variables for cluster classification (distinguish clusters dominated by electromagnetic and hadronic interactions) [125] are exploited. Many of the variables are also the input features in π^0 identification [115]. These variables describe the energy density and shower shape of the topo-cluster of the π^0 candidates. They are added to distinguish the π^0's from τ_{had} decay and those from the other sources. The definition of these variables is outlined in Table 3.4. In the table, the cluster moment of a certain degree n (first, second) for an observable x is defined as

$$\langle x^n \rangle = \frac{1}{E_{\text{norm}}} \sum_{\{i|E_i>0\}} E_i x^n, E_{\text{norm}} = \sum_{\{i|E_i>0\}} E_i, \tag{3.12}$$

where i sums over the ECal cells with positive energy. The distribution of cluster variables of the leading-p_{T} neutral PFOs in different decay modes are shown in Figure 3.4. There are more cluster classification variables, but it is found that some variables are not suitable for the decay mode classification. This is tested by comparing the performance of the DeepSet NN trained with random values assigned to a variable with the DeepSet NN trained with that variable as its original value. Variables are discarded if the performance of the 'random' version is better. The choice of the cluster variables are not further optimised.

In summary, input features consist of kinematics, track impact parameters and cluster properties for τ_{had} tracks (10 variables), conversion tracks (10 variables), neutral PFOs (22 variables) and shot PFOs (6 variables). The values of the variables are standardised by subtracting the means and scaled by the standard derivations. For each type of objects, only the leading n objects are considered, ordered by p_{T}. The maximum number of objects is three, four, eight and six for τ_{had} tracks, conversion tracks, neutral PFOs and shot PFOs, respectively.

Table 3.4: Name and description of the cluster classification variables for neutral PFOs as input features to the DeepSet NN decay mode classification. The ones with log are transformed to the logarithm values ($\log_{10}$) in training and evaluation. More details are described in Ref. [125] and [126].

Name	Description
$\langle\eta\rangle$	First moment in η in shower axis
$\log(\langle r^2\rangle)$	Second moment in r in shower axis
$\Delta\theta$	Distance in θ between the EM shower axis and the vector pointing from the primary vertex to the centre of the shower
$\log(\lambda_{\text{centre}})$	Distance of the shower centre from the calorimeter front face measured along the shower axis
longitudinal	Normalised second longitudinal moment
$\log(\langle\rho^2\rangle)$	Second moment in energy density
f_{core}	Sum of energy fractions in the most energetic cells per sampling
$f_{\text{core}}^{\text{EM1}}$	Same as f_{core} but only consider EM1
$N_{\text{pos,EM1}}$	Number of cells with positive energy in EM1
$N_{\text{pos,EM2}}$	Number of cells with positive energy in EM2
E_{EM1}	Energy in the EM1 layer
E_{EM2}	Energy in the EM2 layer
$\langle\eta_{\text{EM1}}\rangle$ w.r.t. cluster	First moment in η in EM1 with respect to the cluster
$\langle\eta_{\text{EM2}}\rangle$ w.r.t. cluster	First moment in η in EM2 with respect to the cluster
$\log(\langle\eta^2_{\text{EM1}}\rangle)$ w.r.t. cluster	Second moment in η in EM1 with respect to the cluster
$\log(\langle\eta^2_{\text{EM2}}\rangle)$ w.r.t. cluster	Second moment in η in EM2 with respect to the cluster

3.3.2.2 DeepSet neural network

Multiple objects of different types associate with a τ_{had}, which form a graph-like structure. Decision trees are not suitable to extract the full information from such structure because it is almost dedicated to vector-like inputs. Nevertheless, GNNs are designed for this situation. It can extract information and make predictions on the node-, edge- and graph-level. Due to this feature, GNNs are becoming increasingly interested in particle physics community [127].

In the decay mode classification, the information is extracted at node-level (from each object), and the classification is made at graph-level (for each τ_{had}), and there are no edge-level variables involved. Recurrent neural networks might be suitable, but one cannot avoid to define a non-physical order of sequence. If the order changes, RNN may yield different and unpredictable

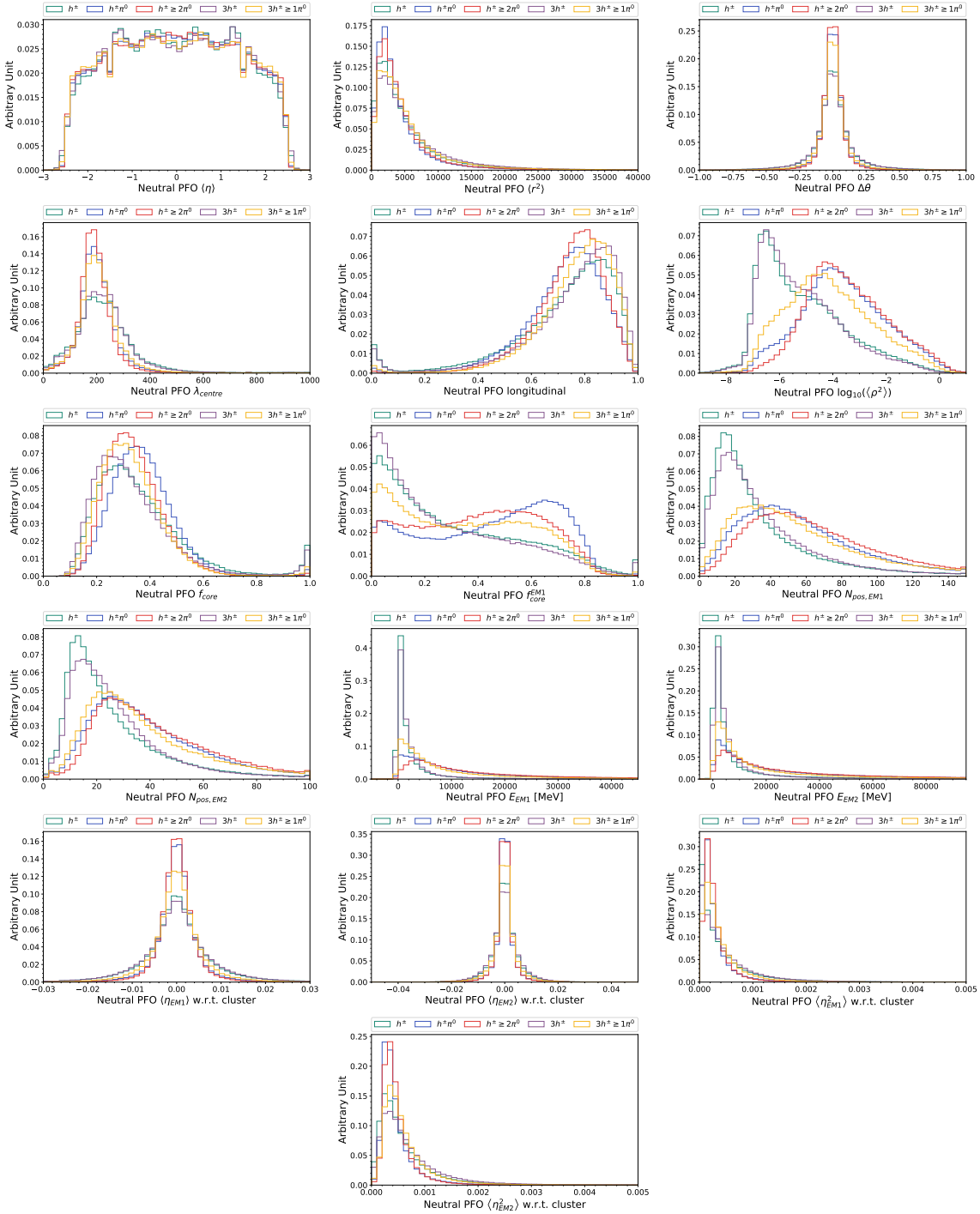


Figure 3.4: Distribution of cluster variables of the leading-p_{T} neutral PFOs in different decay modes, normalised to unit integral.

performances. Meanwhile, RNN is notorious for the difficulties in its training stage. Because each time step depends on the last one, it is slow to train and hard to parallel. DeepSet neural network is a better choice for this kind of problems. It is a special type of GNN, which takes permutation invariant sets (with arbitrary finite size) as inputs and can predict either a regular output or a permutation invariant output. The former is called a invariant model, and the latter is called a equivariant model. Only the invariant model is utilised in the decay mode classification problem. A notable application of DeepSets in ATLAS is the impact parameter based b-jet tagging [128].

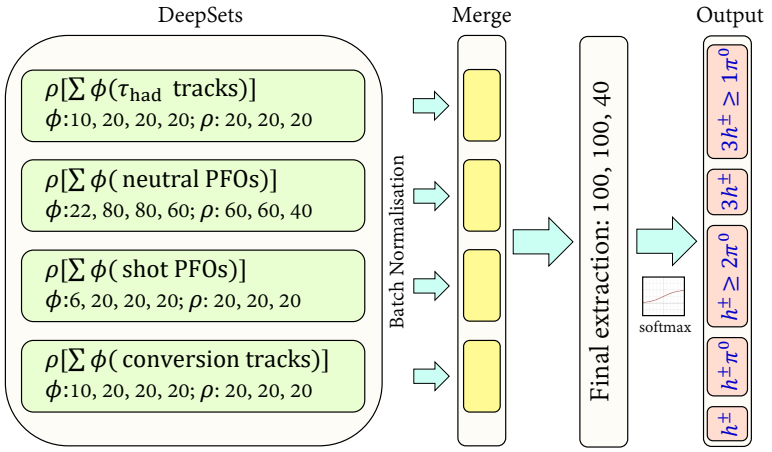
The DeepSet architecture is inspired by the following theorem [121]. Given a function $f(X)$ that operates on a countable set X. The function is invariant to the permutation of elements in X, iff it can be decomposed in the form $\rho\left(\sum_{x\in X}\phi(x)\right)$, for suitable transformations ϕ and ρ. Intuitively, the summation guarantees the permutation invariant property. The $f(X)$ is the decay mode classifier to be learned. In DeepSets, the suitable ϕ and ρ are fitted by training neural networks. The input features $\boldsymbol{x}_i$ of a set of objects $X = \{\boldsymbol{x}_i\}$ are processed by the ϕ NN, outputting $\phi(\boldsymbol{x}_i)$. The outputs are added together and the summation is processed by another neural network ρ. Both ϕ and ρ are feed forward neural networks. For each type of objects, one DeepSet is utilised to process the input features and the output of each DeepSet need to pass a batch normalisation layer to standardise the values in that layer. Lastly, the processed information are merged into one vector to give the final classification after passing several layers of feed forward neural networks.

The implementation and training of the DeepSet are powered by deep learning library Keras [129] with Tensorflow as backend. The architecture and training parameters are summarised in Figure 3.5, where the number of nodes and training parameters are optimised. Only the τ_{had}'s with generated and reconstructed p_{T} greater than 15 GeV, and geometrically matched to the generated-level visible decay products of τ-leptons are used in for the training. The output is a vector of five probabilities corresponding to each decay mode, expressed as 'NN Score (*decay mode*)'. The scores of each generated decay mode are shown in Figure 3.6. The multi-class cross-entropy (CE) loss function (Equation (3.10)) is used. The maximum number of epochs is set to 100, but the training will stop when the loss evaluated on an independent validation sample does not improve in the recent ten epochs . Meanwhile, the learning rate is reduced by a factor of 0.4 when that does not improve in the recent four epochs. The values of the (averaged) CE loss as a function of epoch on training and validation samples are shown in Figure 3.7. No overtraining issues are observed.

3.3.3 Expected performance

The decay mode with the largest probability in the DeepSet output and has a reconstructed number of prong consistent with that based on the τ_{had} track classification [12] is assigned as the reconstructed decay mode. The performance of decay mode classification is evaluated with independent simulated $\gamma^* \to \tau\tau$ samples against the samples for training. The τ_{had} for evaluation need to pass the

[12] Rate of inconsistency is around 1%.



Training Parameters

Parameter	Value
Epochs	100*
Batch size	1000
Optimizer	*Adam*
Learning rate	0.01*
β_1, β_2	0.9, 0.999
Activation function	ReLU

Figure 3.5: Architecture (left) and training parameters (right) of the DeepSet NN decay mode classifier. ρ, ϕ and the layers after merging are feed forward NNs. The numbers indicate the number of nodes in each layer. Details about the epoch and learning are described in the text. *Adam* is a stochastic gradient descent (SGD) optimisation algorithm, which is self adaptive in terms of estimating first and second order moments. β_1 and β_2 are the decay rates for first and second order moment estimates. More information about the parameters can be found in Keras documentation [129].

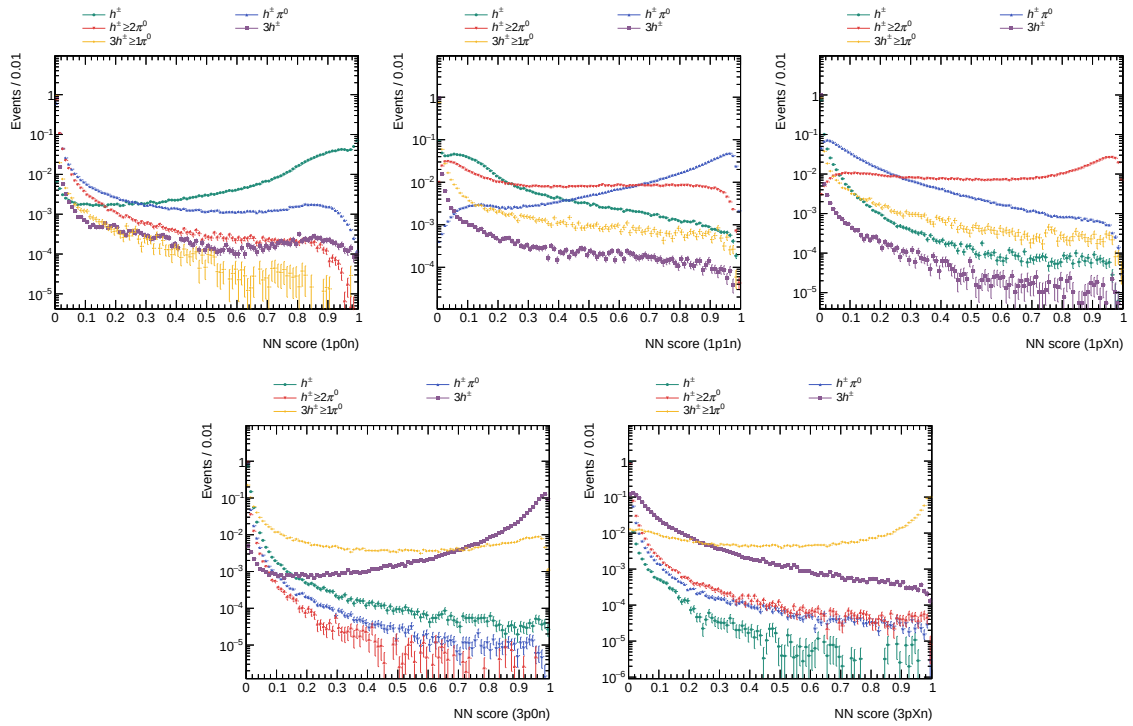


Figure 3.6: Output scores of the DeepSet NN of each generated decay mode. The error bar stands for the statistical uncertainty of simulated $\gamma^* \to \tau\tau$ samples.

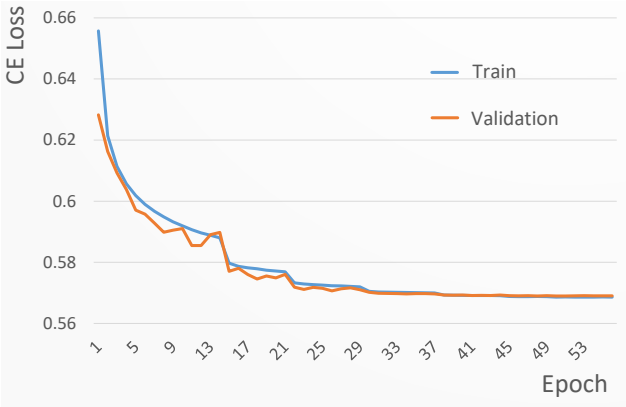
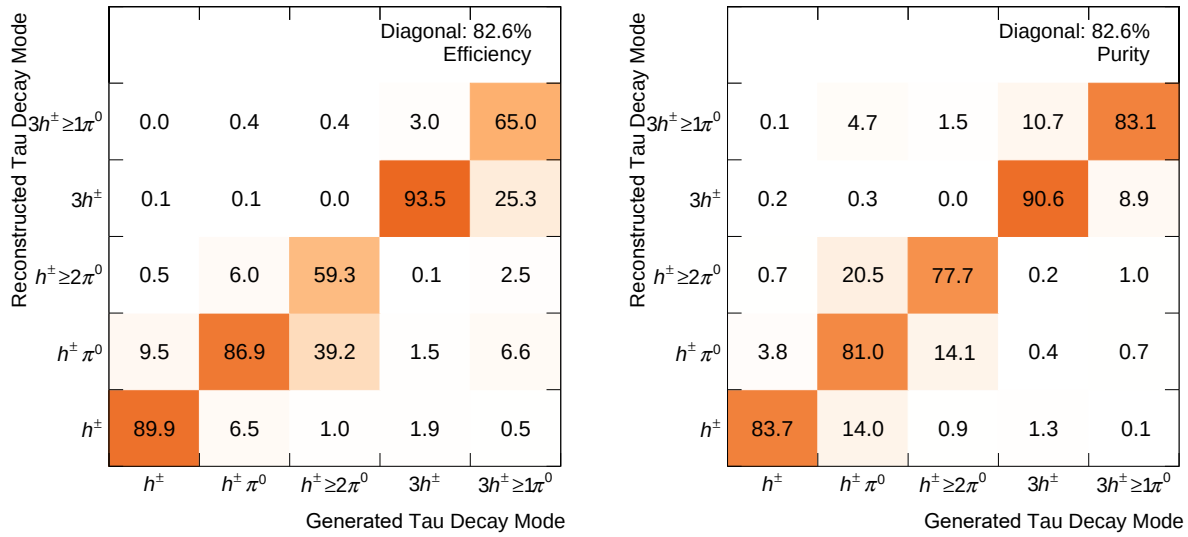


Figure 3.7: Values of the CE loss as a function of epoch on training and validation samples.

p_{T}, η and number of prong (for both generator-level and reconstruction-level variables) requirements in Table 3.2 and the *Medium* τ_{had}-ID working point, which is used in the standard model $H \rightarrow \tau\tau$ physics analysis. The performance of PanTau decay mode classifier is evaluated with the same samples and selections. The efficiency and purity matrices are shown in Figure 3.8. It outperforms the PanTau by 9% on the diagonal efficiency/purity, showing a significant improvement in all decay modes. Figure 3.9 and Figure 3.10 present the improvement as functions of τ_{had} p_{T}, η and average number of bunch crossing (pile-up). The most significant improvements are in the $h^{\pm}\pi^0$ purity and $h^{\pm} \geq 2\pi^0$ efficiency in the lower p_{T} region ($pT < 200$ GeV). In general, the improvements are stable with respect to η and pile-up. Detailed performances especially for the non-diagonal parts of the efficiency/purity matrices can be found in Appendix A.



(a) Efficiency (b) Purity

Figure 3.8: Efficiency (a) and purity (b) matrices of DeepSet NN evaluated on simulated $\gamma^* \rightarrow \tau\tau$ samples.

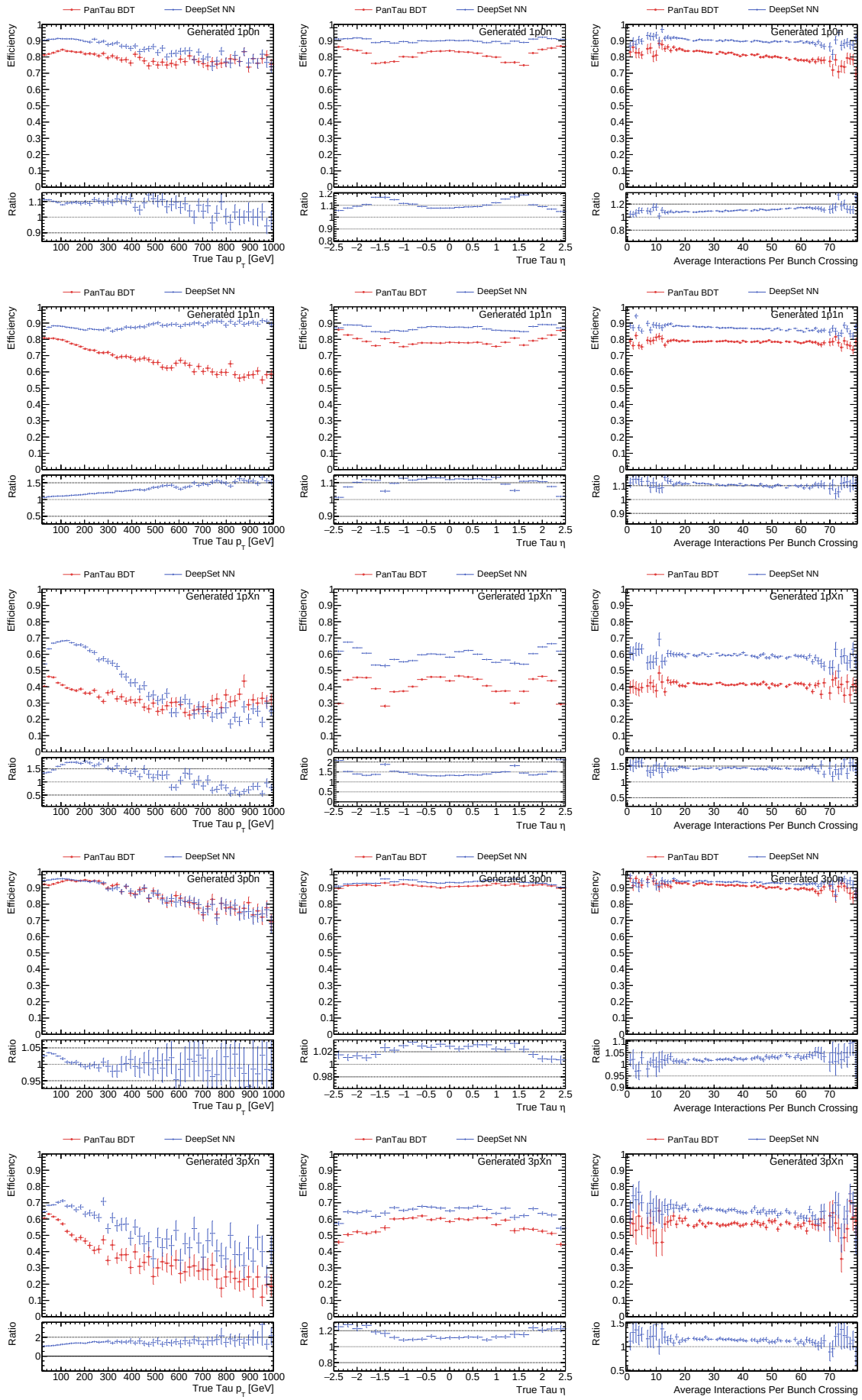


Figure 3.9: Comparison of the decay mode classification efficiency as functions of τ_{had} p_{T}, η and average number of bunch crossing between the PanTau and DeepSet decay mode classifiers, for different generator-level decay modes (in each row). The error bar stands for the statistical uncertainty of simulated $\gamma^* \to \tau\tau$ samples.

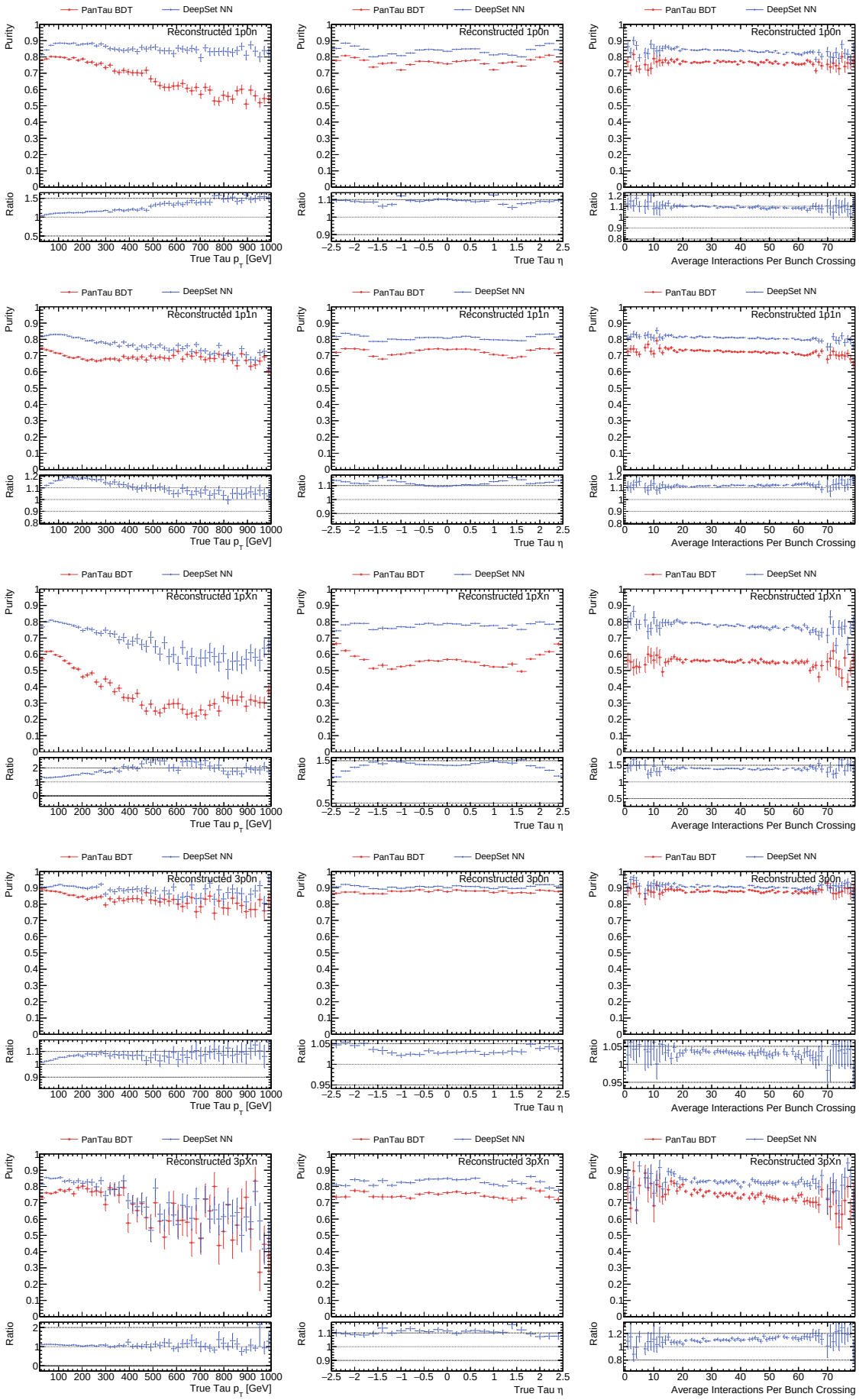


Figure 3.10: Comparison of the purity of classified decay modes as functions of τ_{had} p_{T}, η and average number of bunch crossing between the PanTau and DeepSet decay mode classifiers, for different generator-level decay modes (in each row). The error bar stands for the statistical uncertainty of simulated $\gamma^* \rightarrow \tau\tau$ samples.

3.4 Full-hadronic boosted di-τ object

In the search for resonant $HH \to b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ produced via a heavy scale particle X in Chapter 5, the separation of two τ_{had} from the decay of the Higgs boson become smaller as the X (m_X) is heavier, because the Higgs boson is more Lorentz boosted. It is predicted by the simulated $X \to HH$ samples that if m_X is larger than 2 TeV, the angular separation of two τ_{had} will be smaller than 0.4. Therefore, the standard τ_{had} reconstruction algorithms will become inefficient because it cannot resolve them individually. To recover those events, a novel technique [13] is developed to reconstruct the merged $\tau_{\text{had}}\tau_{\text{had}}$ object, referred to as a full-hadronic boosted di-τ object or di-τ for short in the following content.

Boosted di-τ objects are seeded by untrimmed anti-k_t large-R jets ($R = 1.0$) with a transverse momentum $p_{\text{T}} > 300$ GeV. The constituents of the large-R jets are re-clustered into smaller radius ($R = 0.2$) anti-k_t sub-jets. It is required to have two sub-jets such that the di-τ system can be constructed by the two p_{T}-leading sub-jets. The four-momentum of di-τ is also calculated by the two p_{T}-leading sub-jets. Each sub-jet must have $p_{\text{T}} > 10$ GeV and at least one geometrically matched track that fulfils the same track-vertex matching as the standard τ_{had} track selection and the following track selection

- $p_{\text{T}} > 1$ GeV;
- Number of hits in the pixel detector ≥ 2;
- Number of hits in the pixel detector + in the SCT ≥ 7.

The matched tracks within the $R = 0.2$ cone around the axis of a sub-jet are called 'τ tracks', while those that have $\Delta R \geq 0.2$ but still inside the cone of the seeding large-R jet are called 'iso-tracks'. Figure 3.11 shows the di-τ reconstruction efficiency as a function of the separation $\Delta R(\tau_{1,\text{vis}}, \tau_{2,\text{vis}})$ and p_{T} of the two generator-level visible products of the two τ_{had} (τ_{vis}) that the two leading sub-jets are matched to. The selection before calculating the efficiency on the generator level objects is shown in the figures. The efficiency is estimated with $X \to HH \to b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ simulated samples generated with $m_X = 2$ TeV (see Section 5.1). The standard resolved τ_{had} reconstruction acting on each τ_{had} is compared with the boosted di-τ method. It shows that the boosted di-τ reconstruction recovers the sharp efficiency loss of the standard method when the separation is about $0.2 < \Delta R(\tau_{1,\text{vis}}, \tau_{2,\text{vis}}) < 0.4$ and when p_{T} of di-τ is greater than about 300 GeV.

The energy of the reconstructed di-τ objects is compared with that of the generator-level di-τ_{vis}. They are found to be close, the energy scale as a function of the p_{T} of di-τ_{vis} is given in Figure 3.12(a). The energy resolution as a function of the p_{T} of di-τ_{vis} is shown in Figure 3.12(b). It is found that better resolution is reached at higher energy of di-τ. The performances are evaluated using the same $X \to HH$ ($m_X = 2$ TeV) samples. In Section 5.3.3, it is found that the kinematics of reconstructed di-τ is well modelled by the simulated samples, therefore the (sub-jets of) di-τ is calibrated by the LC weighting scheme mentioned in Section 2.3.3.

[13] Reported in Ref. [130] together with the $HH \to b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ boosted analysis Chapter 5

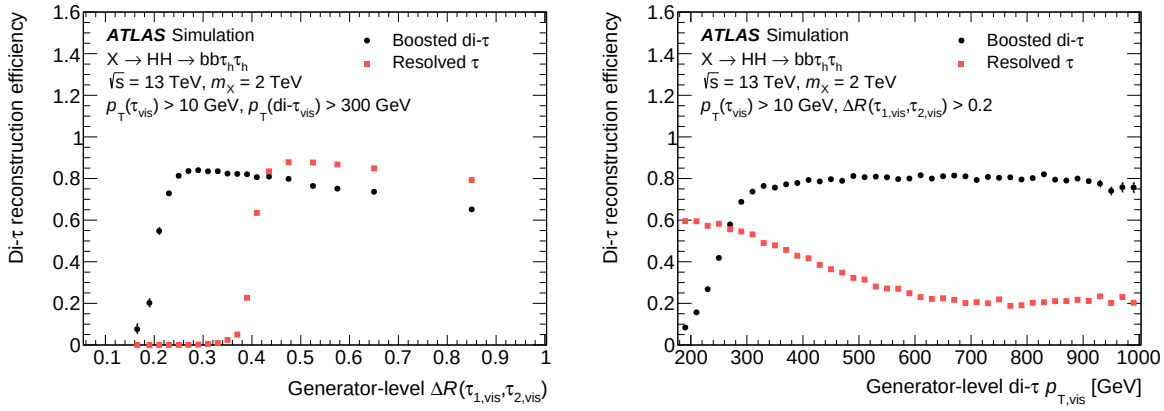
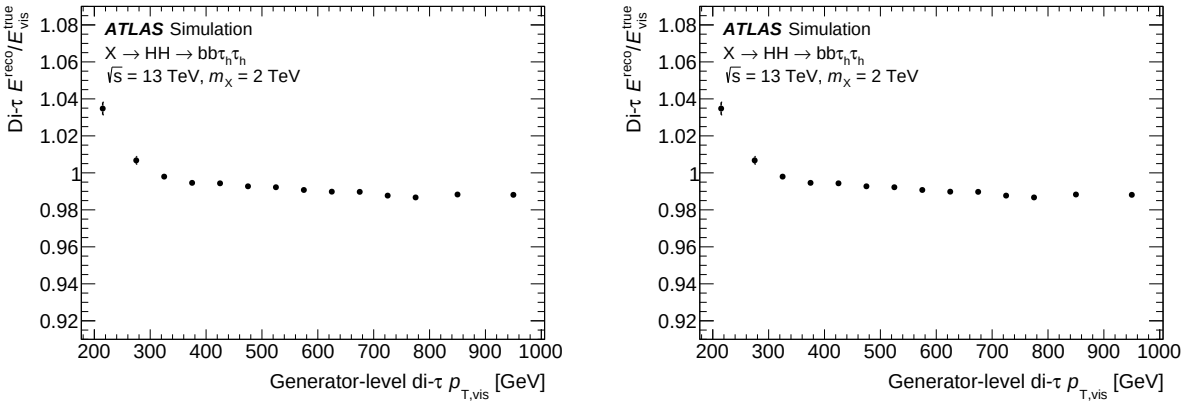


Figure 3.11: Reconstruction efficiency of di-τ system using standard resolved reconstruction in red and boosted reconstruction in black, as a function of the angular distance $\Delta R(\tau_{1,\text{vis}}, \tau_{2,\text{vis}})$ and p_{T} constructed by the matched generator-level τ_{vis} [130]. The statistical uncertainty of the simulated samples is indicated in the vertical error bar.



(a) Energy scale (b) Energy resolution

Figure 3.12: Energy scale (a) and resolution (b) of the di-τ objects as a function of the p_{T} of the generator-level di-τ_{vis} [130]. The degradation of the energy scale at lower di-τ_{vis} p_{T} is caused by the 300 GeV threshold of the seeding large-R jet. The statistical uncertainty of the simulated samples is indicated in the vertical error bar.

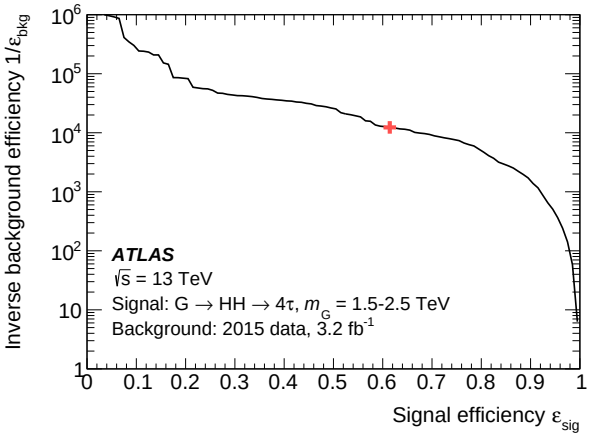


Figure 3.13: di-τ signal efficiency ϵ_{sig} and its corresponding jet background rejection power (inverse background efficiency $1/\epsilon_{\text{bkg}}$) of the di-τ-ID BDT, performing on the reconstructed di-τ [130]. The red cross indicates the 60% efficiency working point which is the baseline di-τ-ID requirement in the boosted analysis.

Reconstructed di-τ is separated against the jets initiated from quarks and gluons. It is the identification of di-τ (di-τ-ID), similar as the τ_{had}-ID for resolved τ_{had}. In di-τ-ID, a BDT variable is trained to separate boosted di-τ from jets based on input features built by the information about tracks, vertices and energy deposits in the calorimeters. More details about the input variables can be found in [130]. In the training of the BDT di-τ-ID, the di-τ objects are simulated from a spin-2 graviton G sample with various mass m_G, where $G \rightarrow HH \rightarrow (\tau\tau)(\tau\tau)$, and the jets are selected from real data recorded in 2015 with an integral luminosity of 3.2 fb^{-1}. The di-τ signal efficiency and its corresponding jet background rejection power (inverse background efficiency) are shown in Figure 3.13. The red cross on the figure represents the 60% efficiency (10^4 rejection) working point of the di-τ identification. It corresponds to a di-τ BDT score greater than 0.72, which is used in Chapter 5 as the baseline di-τ identification selection. The performance of di-τ-ID efficiency is found to be stable for di-τ with $p_{\text{T}} > 300$ GeV and for different number of pile-up interactions.

Finally, same as resolved τ_{had} selection, the number of the matched 'τ tracks' of the leading two sub-jets must be either one or three. The efficiency of this selection at this stage is about 80%, estimated using the same $X \rightarrow HH$ ($m_X = 2$ TeV) samples. Additionally, it can further reduce quark/gluon-initiated jet backgrounds.

Chapter 4

Search for Higgs Boson Pair Production

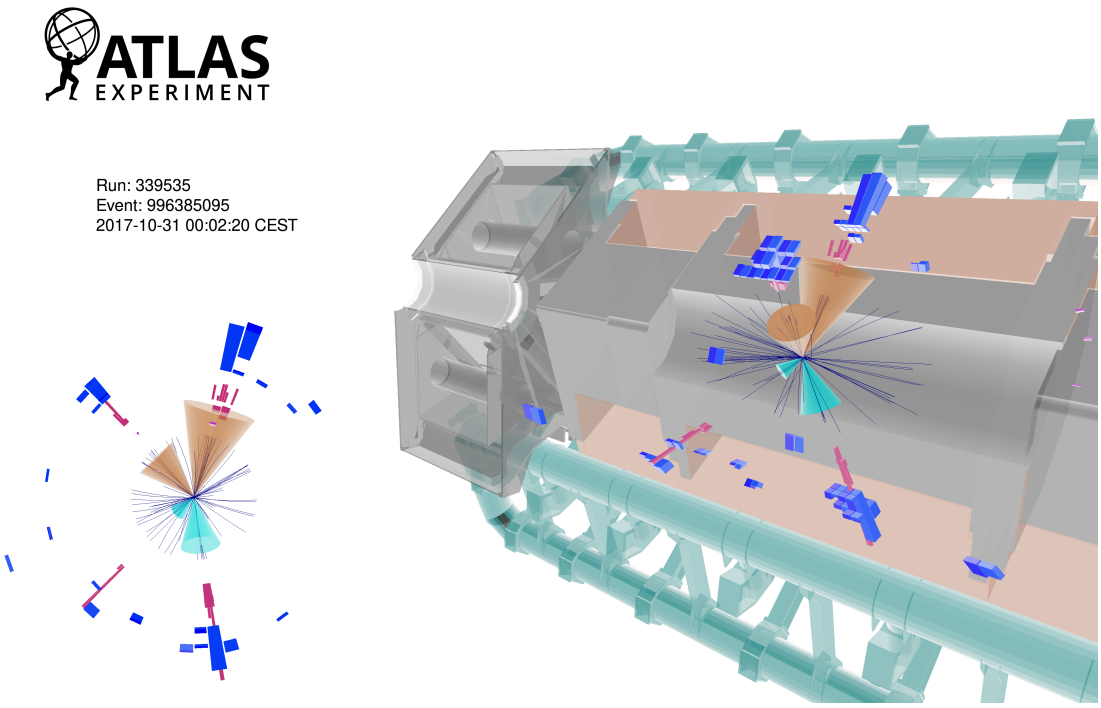


Figure 4.1: Event display of a selected non-resonant $HH \to b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ candidate event.

This chapter elaborates the $HH \to b\bar{b}\tau^+\tau^-$ analysis using the resolved final state topology. This analysis searches both the non-resonant and resonant Higgs boson pair production modes. The fully-hadronic decay channel $\tau_{\text{had}}\tau_{\text{had}}$ is presented in detail (expressed as the $HH \to b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ analysis). This decay channel is combined with the semi-leptonic channel $\tau_{\text{lep}}\tau_{\text{had}}$ into the final results of the $b\bar{b}\tau^+\tau^-$ final state, which is published in [131]. The search for the non-resonant production mode is extended to constrain the Higgs boson self-coupling strength. Related results are published in [132].

4.1 Data and simulated samples

The $HH \to b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ analysis is performed based on the Run-2 proton-proton (p-p) collision data collected by the ATLAS detector (see Section 2.1). Simulated samples (see Section 2.4 for general descriptions) are produced to model the contributions of SM backgrounds and the expectations of HH signals. An overview of the production procedure can be found in Section 2.4. A certain amount of simulated events is generated for each signal and background processes. To be compared with the real data, the simulated events are normalised to the integrated luminosity of Run-2 data and the most precise cross sections by theoretical calculation. A mass of 125 GeV is assumed to the Higgs boson in the simulation. The production and decay of the Higgs boson are according to this value of mass (see Section 1.2.1).

The SM ggF and VBF production modes of HH are considered in the search for non-resonant HH production, as discussed in Section 1.2.2. The ggF samples are generated with full top-quark mass (FT) at NLO by Powheg Box V2 with the PDF4LHC15 [133] PDF set, and the VBF samples is generated by MadGraph5_aMC@NLO at LO with NNPDF3.0nlo [134] PDF set. The parton shower and hadronisation are produced by Pythia 8 [135].

A heavy, narrow-width, scalar particle X that is produced via ggF mode and decays to HH is considered in the search for resonant $X \to HH$ production, as shown in Section 1.2.4. The scalar is assumed to have narrow width (10 MeV) so that the interference between $X \to HH$ and SM HH production can be ignored [6]. The sample is generated at LO by MadGraph5_aMC@NLO using NNPDF2.3lo [123] and the PS generator is Herwig 7 [136]. Several mass hypotheses are generated to cover mass search range from 251 GeV to 1600 GeV. These mass points m_X include

$$251, 260, 280, 300, 325, 350, 400, 450, 500, 550,$$
$$600, 700, 800, 900, 1000, 1100, 1200, 1400, 1600.$$

The unit is GeV. Under m_X =1100 GeV, the samples are simulated with FastSim which is usually valid for BSM searches. At m_X =1100 GeV and above, the FullSim simulated samples are used [1]. The difference between FastSim and FullSim is introduced as a systematic uncertainty.

The SM backgrounds in consideration are W/Z boson produced associated with jets (W+jets, Z+jets), multi-jet production, single top-quark production, top-quark pair production ($t\bar{t}$), top-quark pair produced in association with W/Z ($t\bar{t}W$, $t\bar{t}Z$), diboson production (WW, WZ, ZZ) and single Higgs boson production (with Higgs decays to $b\bar{b}$ or $\tau^+\tau^-$). The major production modes, ggF, VBF, VH and $t\bar{t}H$ samples, are considered in the production of simulated single Higgs boson samples. Multi-jet production is a purely QCD process, whose cross section is overwhelmingly large in the p-p collision. This process tightly suppressed by the τ_{had}-ID, but the contribution to the total background is still not negligible. It is unrealistic to generate simulated sample with the

[1] This is due to technical reason. The high mass FullSim are borrowed from the boosted analysis Chapter 5. The interaction between the τ_{had}'s and the detector material is studied in the boosted analysis therefore the FullSim samples are needed.

required luminosity. Section 4.3.3 describes the estimation of the amount of multi-jet events by a data-driven approach. Apart from that, all the other background processes are modelled by the simulated samples that are centrally produced by ATLAS experiment for all ATLAS physics analysis. The ME and PS generators for the simulated background samples are summarised in Table 4.1.

Table 4.1: Generators for the simulated background samples.

Process	ME generator (PDF set)	PS generator
W/Z+jets	SHERPA 2.2 [137] (NNPDF3.0NNLO)	SHERPA 2.2
Top-quark	POWHEG BOX V2 (NNPDF3.0NLO)	PYTHIA 8
Diboson	SHERPA 2.2	SHERPA 2.2
Single-Higgs	POWHEG BOX V2 (NNPDF3.0NLO)	PYTHIA 8

4.2 Signal region selections

To extract the HH signal events from the background events, a signal-enriched region called signal region (SR) is defined by applying a sequence of selections from the trigger-level to object- and event-level. Finally, multivariate analysis techniques are used to further discriminate HH events and background events.

4.2.1 Trigger selections

The trigger system is introduced in Section 2.2.4. The L1+HLT τ_{had} triggers are used according to the signature of this decay channel. On the basis of the L1 seed triggers, the Single-τ_{had} (STT) and the Di-τ_{had} HLT Triggers (DTT) with the lowest p_{T} threshold are utilised [138]. To sustain sufficiently low event triggering rates (trigger rates), the p_{T} thresholds are tuned and some additional objects are considered. In the HLT, the τ_{had} candidates are selection in three stages [138]:

- Perform a calo-only reconstruction and a dedicated tau energy scale and then selection τ_{had} candidates with a minimum p_{T}.
- Add track information to τ_{had} using the two-stage fast tracking [61] method. Select the τ_{had} candidates with $1 \leq N_{\text{core}}^{\text{trk}} \leq 3$ and $N_{\text{iso}}^{\text{trk}} \geq 1$, where $N_{\text{core}}^{\text{trk}}$ and $N_{\text{iso}}^{\text{trk}}$ are the track multiplicities in $\Delta R < 0.2$ and $0.2 < \Delta R < 0.4$ around the τ_{had} direction.
- Applying offline-like τ_{had}-ID selection, the *Medium* working point is chosen.

After each stage of selections, the trigger rates are reduced so that the next stage is able to be executed. At the same time, events with real τ_{had}'s are kept as much as possible. Offline selections are applied based on the HLT selections. The offline selection criteria are tighter than those of the HLT to insure that only the fully efficient region (i.e. plateau of the trigger efficiency turn-on curve) is used. The offline τ_{had}'s are requirement to be matched to the HLT τ_{had}'s that triggered the

event. This is done by a geometrical matching with $\Delta R < 0.2$. The τ_{had}'s that pass this requirement are called trigger-matched τ_{had}'s. The STT and DTT trigger selections are described below.

The STT triggers require at least one τ_{had} candidate with a minimum p_{T} of 80, 125, 160 GeV (depending on the data-taking period) and pass the track and *Medium* BDT-based τ_{had}-ID selection as mentioned above, namely `medium1_tracktwo`. The offline p_{T} thresholds are 20 GeV higher than the online thresholds. In 2018, a tighter selection is applied on the track to improve the efficiency of selecting 3-prong τ_{had}'s. This set of selections is named as `medium1_tracktwoEF`. After a technical stop during the 2018 data-taking (i.e. starting from period K of 2018 data-taking), MVA-based energy scale and RNN-based τ_{had}-ID superseded the prior algorithms. This new set of selections is named as `mediumRNN_tracktwoMVA`. A logical `or` between the triggers using `medium1_tracktwoEF` and `mediumRNN_tracktwoMVA` selections are used in this period.

The DTT triggers require at least two τ_{had} candidates with minimum p_{T} of 35 GeV and 25 GeV. The offline p_{T} thresholds are 40 GeV and 30 GeV, respectively. For each of the τ_{had}'s in DTT, the configuration of selections in 2018 is the same as that for the τ_{had}'s in STT. Additional offline selections need to be applied to the events triggered by DTT, because extra objects and topologies between objects are required in DTT. They are summarised as follows:

- For events in 2015–2016, an additional L1 jet with minimum p_{T} of 25 GeV is attached to the DTT triggers. An offline selection of $p_{\text{T}} > 80$ GeV is applied to the p_{T}-leading jet, to minimise the inconsistency of trigger efficiencies in data and simulated samples.
- For events in 2017–2018, two categories of additional requirements are used, *4J12* and *L1Topo*. The *4J12* requires two additional L1 jets with minimum p_{T} of 12 GeV. In 2018, the two additional jets are further required to be have a $|\eta|$ smaller than 2.3. The *L1Topo* requires the same jet as in 2015–2016 but also a ΔR between two τ_{had} smaller than 2.8.
- For an event triggered by DTT, if the p_{T}'s of the 1$^{\text{st}}$ and 2$^{\text{nd}}$ p_{T}-leading offline jets are greater than 45 GeV, then the event is categorised into the *4J12*. If an event is not categorised into *4J12* but the p_{T}-leading jet has $p_{\text{T}} > 80$ GeV and the $\Delta R(\tau_{\text{had}}, \tau_{\text{had}}) \leq 2.5$, then it is categorised into the *L1Topo* [2].

The efficiencies of the trigger selections are measured in both data and simulated samples. The efficiencies of simulated samples are calibrated to data samples by trigger scale factors [138, 139].

4.2.2 Event selections

Events that survive the trigger selections enter the stage of offline event selection. Event-level variables are constructed from the physics objects that are selected in Section 2.3 and for τ_{had} in Section 3.2. The events are required to have no electrons or muons, exactly two τ_{had}'s with *Loose* τ_{had}-ID quality and opposite-sign of electric charges (OS), and at least two jets of which exactly two are tagged as b-jet by the *DL1r* algorithm operating at 77% b-tagging efficiency. The

[2] e.g. The two DTT categories are orthogonal.

selections that correspond to the trigger selections (see Section 4.2.1) are applied at this stage. The mass of the di-τ system estimated by the Missing Mass Calculator (MMC) [140] is required to be larger than 60 GeV to suppress the low mass Drell-Yan process. Due to the neutrinos in the decays of the τ-leptons, the invariant mass of di-τ system cannot be fully reconstructed. The MMC defines additional kinematic constraints by a likelihood approach to estimate the most probable mass $m_{\tau\tau}^{\text{MMC}}$.

4.2.3 Summary of the signal region selections

The SR selections are summarised in Table 4.2. The table is separated into `STT` and `DTT` columns to show the different selections. For the events that pass both the `STT` and `DTT` triggers (with τ_{had}'s trigger-matched), the offline selections of `STT` are applied. There is always a trade-off between selecting signal events more efficiently and rejecting background events more tightly. At this stage of event selection, the goal is to keep as much signal events as possible while suppressing a sufficient amount of background events. The strategy is to let the MVA models perform the more stringent discriminations.

The analysis acceptance times selection efficiency (expressed as $\mathscr{A} \times \epsilon$) of signal events is mainly limited by the trigger selections, the reconstruction and identification of τ_{had}'s, and the tagging of b-jets. The cumulative $\mathscr{A} \times \epsilon$ at each stage of the selection chain is estimated on the simulated non-resonant and resonant signal samples in Table 4.3. After applying all the selections, the $\mathscr{A} \times \epsilon$ for ggF+VBF non-resonant signal events is 4.0%, while for resonant signal events, it varies from 0.61% to 10% depending on the mass, as shown in Figure 4.2.

The SR selections can be altered to define regions that are orthogonal to the SR but contain events with similar signatures as those in the SR. These regions are usually designed to be enriched in a specific background process and free of signal events, such that they can be used to estimate the contribution of that background (in this case the regions are called Control Regions, CRs) or to validate the goodness of the estimation of that background (in this case the regions are called Validation Regions, VRs). Section 4.3 is dedicated to the background estimation, where the definition of the CRs/VRs for background studies will follow.

4.2.4 Multivariate analysis

The MVA models are employed to discriminate signal processes from background processes and maximise the sensitivity of the search. It is a supervised binary classification problem in machine learning, where the output is a scalar score that measures the likelihood of the event being a signal event (see Section 3.1).

Table 4.2: Summary of object-level and event-level selections that defines the SR of $HH \to b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ analysis. $N(X)$ denotes the number of object X. The values of p_{T} that are separated by commas indicates that they correspond to the online threshold. The p_{T} selection of the leading (sub-leading) p_{T} objects are referred to as $p_{\text{T}}^{\text{lead}}$ ($p_{\text{T}}^{\text{subl}}$).

Variable	STT	DTT
τ_{had}		
$p_{\text{T}}^{\text{lead}}$	$> 100, 140, 180$ GeV	> 40 GeV
$p_{\text{T}}^{\text{subl}}$	> 25 GeV	> 30 GeV
Jet		
$p_{\text{T}}^{\text{lead}}$	> 45 GeV	Trigger dependent
$p_{\text{T}}^{\text{subl}}$	> 20 GeV	(see Section 4.2.1)
Event-level		
Trigger	Pass and τ_{had} trigger-matched	
$N(e/\mu)$	$= 0$	
$N(\tau_{\text{had}})$	$= 2$, with opposite-sign electric charge	
$N(\text{jets})$	≥ 2, exactly two b-tagged	
$m_{\tau\tau}^{\text{MMC}}$	> 60 GeV	

Table 4.3: Cumulative analysis acceptance times selection efficiency at each stage of the selection chain estimated on the simulated non-resonant (ggF and VBF) and resonant ($X(m_X$ in GeV)) signal samples. The efficiencies are calculated with respect to $HH \to b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ events. Due to technical reason the selections are not applied in the natural way. The 'Pre-selection' contains a loose kinematic selection on τ_{had}, at least one τ_{had} to pass *Loose* τ_{had}-ID requirement, the veto of electrons and muons and other basic selections. The 'Object selections' requires exactly two τ_{had} and at least two jets with the kinematic requirements mentioned in previous chapters (2.3.3 and 3.2). The 'DTT offline jet selections' refers to those mentioned in Section 4.2.1.

Selection stage	ggF	VBF	$X(300)$	$X(500)$	$X(1000)$
Pre-selection	31%	26%	22%	32%	41%
Trigger	13%	10%	6.5%	15%	29%
Object selections	12%	8.9%	6.0%	15%	29%
Loose τ_{had}-ID	10%	7.3%	4.7%	12%	23%
Opposite-sign of τ_{had} charges	9.9%	7.1%	4.7%	12%	23%
Two b-tagging jets	4.6%	2.9%	1.9%	5.3%	11%
Sub-leading τ_{had} $p_{\text{T}} > 25$ GeV	4.5%	2.9%	1.9%	5.3%	11%
At least one jet with $p_{\text{T}} > 45$ GeV	4.4%	2.7%	1.7%	5.2%	11%
DTT offline jet selections	4.2%	2.3%	1.1%	5.2%	11%
$m_{\tau\tau}^{\text{MMC}} > 60$ GeV	4.1%	2.3%	1.1%	5.1%	10%

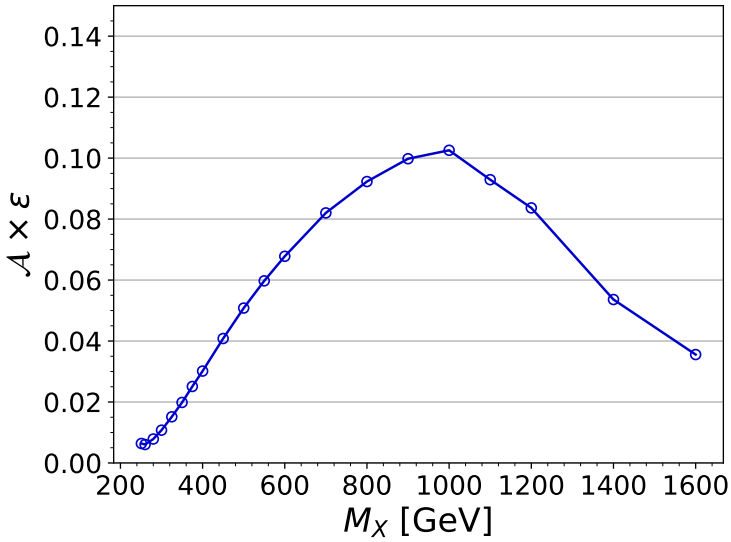


Figure 4.2: Analysis acceptance times selection efficiency of the full signal region selections simulated with the simulated resonant signal samples as a function of the mass of the resonant particle m_X.

4.2.4.1 Input variables

Input variables and targets are needed in the supervised learning problem. The targets are apparently known, which are true signal or background by the way that the simulated samples are produced. The input variables are constructed by the distinctive signatures of the signal processes. Five variables are chosen in this analysis;

- m_{HH}: the invariant mass of the HH system constructed from the four-momentum of di-τ_{had} (using MMC) and four-momentum of di-b-jet;
- m_{bb}: the invariant mass of the di-b-jet system;
- $m_{\tau\tau}^{\text{MMC}}$: the invariant mass of the di-τ_{had} system estimated by MMC;
- $\Delta R(b, b)$: angular distance between two b-jet;
- $\Delta R(\tau_{\text{had}}, \tau_{\text{had}})$: angular distance between two τ_{had}.

They provide good separation between signals and backgrounds, as illustrated in Figure 4.3. No clear linear correlation patterns are found between the variables, thus these variables are directly used without extra transformations. For the neural networks (see Section 4.2.4.2), the input variables are normalised by subtracting medians and dividing by interquartile ranges [3] for numerical stability.

4.2.4.2 Configuration of the MVA models

In the search for non-resonant HH signal, the BDT model (implemented by TMVA) is chosen as the classifier. The neural network based method is also checked and it shows almost identical

[3] The medians and interquartile ranges are derived on the simulated signal+background samples.

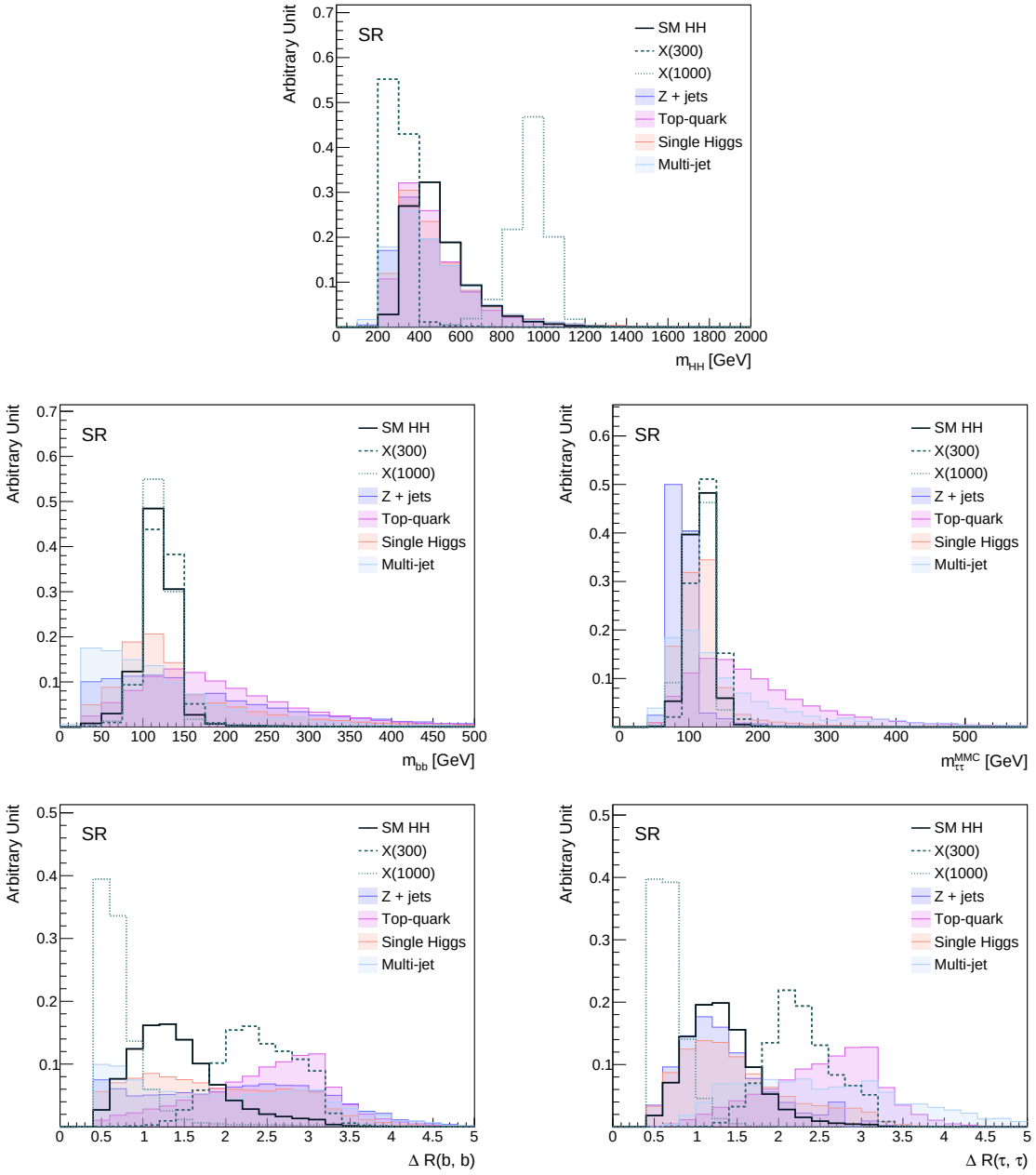


Figure 4.3: Distribution of the MVA input variables of various signals and the most important backgrounds normalised to identity. The $X(m_X)$ denotes the resonance signal generated with mass m_X.

performance as the BDT-based one. The configurations for the BDT training are listed in Table 4.4, where the parameters are found to be optimal.

Table 4.4: The parameters of the non-resonant HH BDT model. Definitions of the parameters can be found in [112].

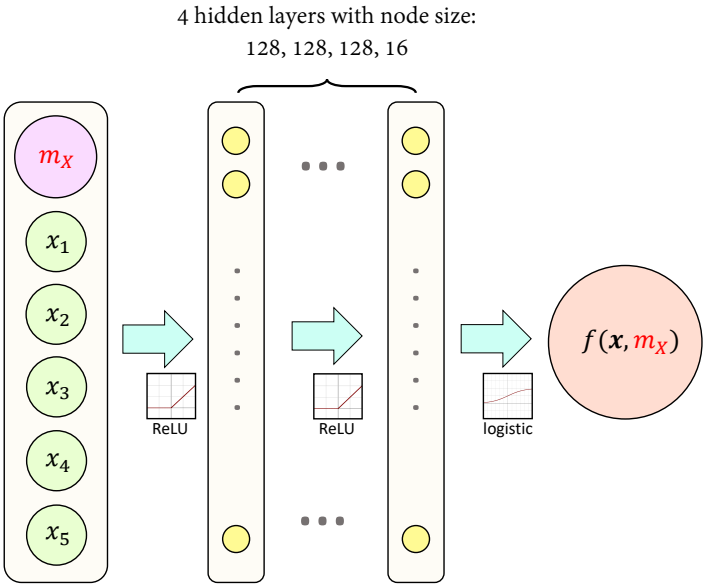
Parameter	Value
`NTrees`	1500
`MaxDepth`	2
`MinNodeSize`	1%
`BoostType`	*GradientBoost*
`Shrinkage`	0.2
`IgnoreNegWeightInTraining`	Yes

In the resonant $X \rightarrow HH$ case, the performance needs to be optimised for a wide range of resonance mass (251–1600 GeV). Traditionally, different models are trained for difference signal mass hypothesis m_X. To simplify the training procedure, a new structure called Parameterized Neural Network (PNN) [141] is utilised. Compared with the plain NN models, the PNN adds one or more macroscopic parameters of a classification problem to the list of input features, such that the output of the PNN is a function of those parameters and the performance of the PNN is optimal for those parameters. Thus, the optimisation can be accomplished in a single parameterized model. Another advantage is that the optimisation is not only for the parameter values that appear in the input, but can also be generalised to the interpolated values smoothly. This is also the reason that BDT is not suitable for this problem , because it is a discrete algorithm. Here, the parameters is the mass hypothesis of the resonance particle (m_X). In practise, signals of all mass points are used in the training and they are rescaled to have equal sum of weights to create a (discrete) uniform distribution of m_X for the training. The backgrounds do not have definition of m_X therefore it is randomly assigned according to the same distribution as the signal. The architecture of the PNN is sketched in Figure 4.4, in which the optimal configurations of the PNN training are indicated. The training of the PNN is powered by Keras with Tensorflow as backend.

In both cases, the weights of the simulated samples are considered in the training and events with negative weights are ignored. The signal and background events are scaled to the same sum of weights in order to have balanced classes for training. The full dataset is randomly split into two sets and trained independently. In the evaluation stage, the predictions of the events in one set are made by the model trained by the events in the other set. For real data, one of the models is used randomly.

4.2.4.3 Output

The non-resonant BDT score and the resonant PNN scores (together expressed as MVA scores) with m_X = 300, 1000, 1600 GeV (labelled as PNN300, etc.) of the signals and the major backgrounds are depicted in Figure 4.5. The MVA models show powerful separation of signal and background processes. Discriminations of the single Higgs boson and Z+jets processes are worse



Training Parameters

Parameter	Value
Epochs	200
Batch size	128
Optimizer	*SGD*
Learning rate	0.2
Learning rate decay	1e-5
Nesterov momentum	0.9

Figure 4.4: Illustration of the architecture (left) and training parameters (right) of the PNN, where x_i ($i = 1...5$) are original input variables (m_{HH}, m_{bb}, $m_{\tau\tau}^{\text{MMC}}$, $\Delta R(b,b)$, $\Delta R(\tau_{\text{had}}, \tau_{\text{had}})$), and m_X is the resonance mass parameter of the PNN. It is a feed forward neural network. ReLU and logistic are the activation functions. The logistic function provides a parameterized scalar score $f(\mathbf{x}, m_X)$ which range from 0 to 1. The score measures the possibility of the event of being a signal event. *SGD* stands for the optimizer using a stochastic gradient decent algorithm. More information about the parameters can be found in Keras documentation [129].

than the that of other processes. This is because these processes are irreducible [4] and the kinematics signature of these processes is very similar to those of the signal processes. It is found that the MVA score distributions on training and testing samples are compatible. Therefore the models are not overtrained.

4.2.4.4 Importance ranking of input features

Importance ranking of the input features is checked to understand the relative contribution of each variable to the discrimination. The ranking is estimated by the relative sensitivity loss ($\Delta Z/Z$) caused by the removal of that feature in the training of the MVA model. The sensitivity is measured by a combined asymptotic significance [142] calculated based on the optimal binning of the MVA score. The combined asymptotic significance is

$$Z = \sqrt{\sum_k z_k^2}, \tag{4.1}$$

where

$$z_k = \sqrt{2(s_k + b_k)\ln\left(1 + \frac{s_k}{b_k}\right) - 2s_k} \tag{4.2}$$

[4] For single Higgs boson, the 'irreducible processes' refer to VH and ttH.

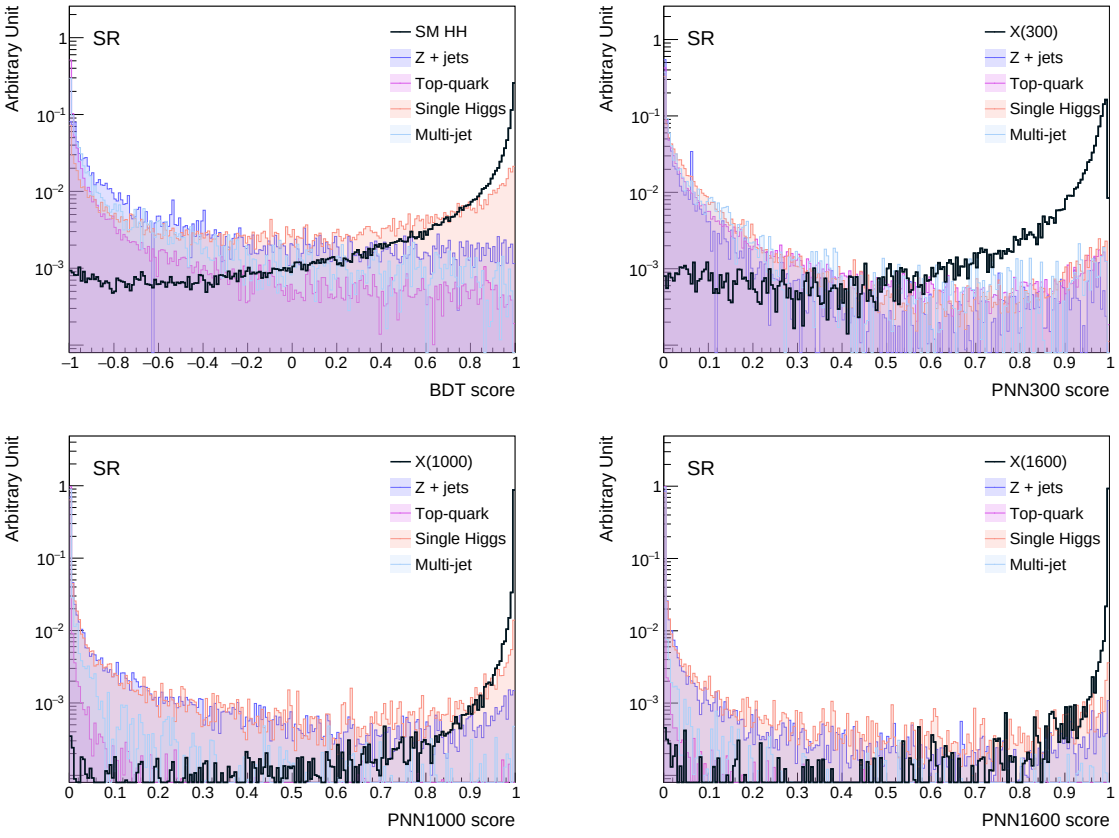


Figure 4.5: MVA scores of the signals and most important backgrounds – BDT score for non-resonant HH and PNN score for resonant HH. The integral of the distributions are normalised to one.

is the asymptotic significance of the k-th bin, and s_k and b_k are the number of signal and background events in the k-th bin, respectively. The optimal binning is represented by a sequence of MVA scores $\{e_i\}$ which form the bin edges. The bin edges are selected from another sequence of MVA scores $\{c_i\}$, called the bin edge candidates. The binning is decided by the following routine:

- The highest MVA score is defined as the first bin edge candidate c_0 and the first bin edge e_0, while the lowest score is defined as the last bin edge e_{last}.
- The MVA score is scanned from highest to lowest value with a stride of $\frac{e_0 - e_{\text{last}}}{1000}$. If the signal and background events in the bin between the current score and the latest bin edge candidate c_j satisfy $n_b > 5$ and $\sigma_b < \min(20\%, \epsilon \frac{n_s}{n_s^{\text{total}}})$, then the current score is a new bin edge candidate c_{j+1} [5]. Here n_b, σ_b are the number and statistical uncertainty of the background events in that bin. The n_s is the number of signal events in that bin and n_s^{total} is the total number of signal events.
- Once the scan is finished, the bin edge candidates $\{c_i\}$ are found. A significance Z_{max} can be calculated based on $\{c_i\}$, it is approximately the maximum expected significance.
- Starting from a single bin, which can be labelled as $\{e_{\text{last}}, e_0\}$, choose c_j from $\{c_i\}$ that gives the largest Z under the binning of $\{e_{\text{last}}, c_j, e_0\}$ and add c_j to the bin edges. Next, choose c_k from $\{c_i | i \neq j\}$ that gives the largest Z under the binning of $\{e_{\text{last}}, c_k, c_j, e_0\}$ [6] and add c_k to the bin edges.
- Bin edges are added repeatedly until $1 - \frac{Z}{Z_{\text{max}}} < 10^{-5}$, i.e., the significance does not improve by adding new bin edges. When the routine stops, the optimal binning $\{e_i\}$ is found.

The requirement $n_b > 5$ is applied in order to use the asymptotic formula [142]. The purpose of $\sigma_b < \min(20\%, \epsilon \cdot \frac{n_s}{n_s^{\text{total}}})$ is to insure sufficiently large size of background samples and to control the number of bins with very few signals (by the parameter ϵ that is set to 1 here). The binning is then transformed to have equal bin width between the maximum and minimum scores for better visualisation in the figures that follow, such as in Figure 4.25.

The rankings of the input features in various MVAs are summarised in Figure 4.6. The input features are of roughly equal importance in the non-resonant HH BDT. The mass variables stand out in the resonant $X \to HH$ PNNs, in particular, the m_{HH} contributes the major discrimination in the high mass PNNs.

4.3 Background estimation

Many background processes can produce $b\bar{b}\tau^+\tau^-$ in their final state. The largest contributions to the background are from $Z \to \tau\tau$ produced in association with heavy flavour (hf) jets ($Z \to \tau\tau + hf$), top-quark pair with/without generator-level true τ_{had}'s (called true/fake-τ_{had} $t\bar{t}$, respectively)

[5] The bin edge candidates satisfy $c_{j+1} < c_j$

[6] The binning can also be $\{e_{\text{last}}, c_j, c_k, e_0\}$

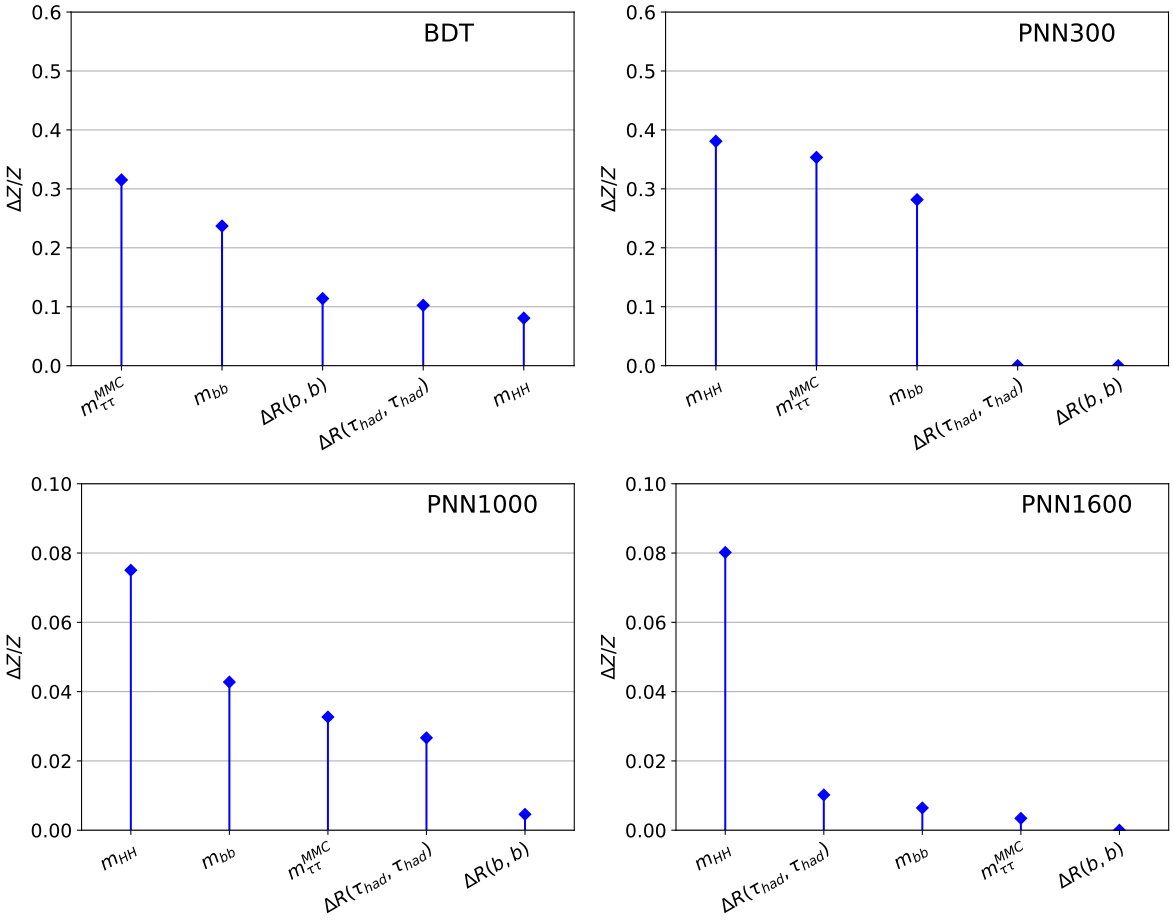


Figure 4.6: The importance ranking for the MVA input features. Features with larger $\Delta Z/Z$ are more important.

and multi-jet. Estimations of these background processes are explained thoroughly below and the validations are summarised in Appendix B.3. The relatively minor processes are directly estimated by simulated samples. These processes include $Z \to \tau\tau$ produced in association with light flavour (lf) jets ($Z \to \tau\tau + lf$), $Z \to \ell\ell$ + jets ($\ell = e/\mu$), $W \to \tau\nu$ + jets, $W \to \ell\nu$ + jets, single top-quark, diboson, $t\bar{t}W/Z$ and single Higgs boson. Table 4.5 lists the pre-fit event yields of data, expected background and SM HH processes after the SR selections.

Table 4.5: Number of data, expected background and SM HH events in the SR. The uncertainty includes all the theoretical, experimental and statistical uncertainties.

Process	Yield ±Uncertainty
$Z \to \tau\tau + hf$	1200 ±210
True-τ_{had} $t\bar{t}$	3600 ±310
Fake-τ_{had} $t\bar{t}$	2500 ±330
Multi-jet	1400 ±150
Single Higgs boson	74 ±15
Other backgrounds	500 ±96
Total background	9200 ±640
Data	8400
SM ggF HH	5.4 ±0.92
SM VBF HH	0.17 ±0.014

4.3.1 $Z \to \tau\tau$ + heavy flavour jets

THe $Z \to \tau\tau + hf$ process is one of the largest background components in this analysis, and it is irreducible. Vector boson + heavy flavour jets is one of the most interesting and challenging processes to be modelled, because the associated heavy flavour jets can be initiated from the sea quark of the incoming proton or from a gluon splitting into a pair of heavy flavour quarks [143]. The flavour (heavy or light) is labelled according to a geometrical matching between the generator-level hadrons and the b-jet candidates. If a b-hadron is matched then label the jet as b-jet. Otherwise if a c-hadron is matched then label the jet as c-jet. Finally, if neither b- or c-hadrons are matched then label the jet as l-jet (light flavour). After applying this procedure, $Z \to \tau\tau$ + jets can be split into $Z \to \tau\tau + bb, bc, cc, bl, cl, ll$. In this analysis, we define heavy flavour as $hf = bb, bc, cc$ and light flavour as $lf = bl, cl, ll$ [7].

There is a known under-estimation of $Z + hf$ events by the SHERPA 2.2 simulated sample as reported by the $V(\to 2\ell)H(\to b\bar{b})$ ($\ell = e/\mu$) analysis [81]. A $Z + hf$ control region (CR) targeting $b\bar{b}\ell^+\ell^-$ final state is defined to derive the normalisation correction factor for $Z + hf$ process. The $b\bar{b}\ell^+\ell^-$ final state can be used because the production of jets is independent of the decay of the Z boson. At the same time, the size of a $b\bar{b}\ell^+\ell^-$ CR is much larger than that of a $b\bar{b}\tau^+\tau^-$-based CR

[7] In the future, we may consider to move bl to hf category as the boosted analysis (see Chapter 5). Nevertheless bl is only 4% of the total $Z \to \tau\tau$ so the impact of categorisation is expected to be negligible.

because of the excellent efficiency and purity of electron and muon selections. The $Z + hf$ CR is defined by the following criteria:

- Pass single electron/muon (SET, SMT) or di-electron/muon (DET, DMT) or electron+muon (EMT) triggers [144, 145]. Online and offline p_{T} thresholds and listed in Table 4.6;
- Exactly two electrons or muons with same flavour and opposite-sign charges;
- Exactly two b-tagged jets;
- Invariant mass of di-lepton system $m_{\ell\ell}$ ($\ell = e, \mu$) is close to the Z boson mass, $75 < m_{\ell\ell} < 110$ GeV;
- $m_{bb} < 40$ GeV or $m_{bb} > 210$ GeV (to reject potential contamination of $HH \to b\bar{b}\ell^+\ell^-$ signal events).

Table 4.6: The p_{T} thresholds of the lepton triggers employed in $Z + hf$ CR. The online and offline p_{T} thresholds are displayed outside and inside the brackets and the pairs are separated by commas. The p_{T} threshold depends on the data-taking period. The form 'A + B' indicates the $p_{\mathrm{T}}^{\mathrm{lead}} + p_{\mathrm{T}}^{\mathrm{subl}}$ in the case of di-lepton triggers and $p_{\mathrm{T}}^{e} + p_{\mathrm{T}}^{\mu}$ in the case of electron + muon triggers.

Trigger	Online (Offline) p_{T} threshold [GeV]
SET	24 (25), 26 (27), 60 (61), 120 (121), 140 (141)
SMT	20 (21), 26 (28), 50 (53)
DET	12 (13) + 12 (13), 17 (18) + 17 (18), 24 (25) + 24 (25)
DMT	18 (19) + 8 (10), 22 (24) + 8 (10)
EMT	17 (18) + 14 (15), 26 (27) + 8 (9), 7 (9) + 24 (26)

The normalisations of $Z + hf$ (μ_{Z+hf}), $t\bar{t}$ ($\mu_{t\bar{t}}$) and the signal (signal strength, μ) are free parameters in the (binned) profile likelihood model. They are determined by the maximum likelihood estimation (a.k.a likelihood fit) in the SR + $Z + hf$ CR joint model (see Section 4.5.1). The model is built based on the BDT/PNN score (SR) and $m_{\ell\ell}$ ($Z + hf$ CR) distributions. In the likelihood fit of the model under background-only hypothesis ($\mu = 0$) using the BDT score template, the best-fit values and 1σ confidence intervals are $\mu_{Zhf} = 1.4 \pm 0.13$ and $\mu_{t\bar{t}} = 0.96 \pm 0.042$. Figure 4.7 illustrates the data–background comparison of the $m_{\ell\ell}$ distribution after applying the above normalisation factor in the $Z + hf$ CR, where good agreement is achieved [8].

As shown later in Section 4.6, the $Z + hf$ process is the most important background in the most sensitive BDT and PNN score bins. Therefore, it is crucial to validate the BDT and PNN responses of the $Z + hf$ process. This is performed by emulate the MVA inference in the $Z + hf$ CR, where the input variables that are constructed with τ_{had}'s are replaced by the ones that are constructed by the electrons or muons. A validation region is designed to be enriched in $Z + hf$ events and to be close to the signal region, which is called a $b\bar{b}\ell^+\ell^-$ $Z + hf$ VR. This VR is defined by closely following the $Z + hf$ CR, but with some different selections listed below.

[8] The experimental uncertainty in the figures are explained in 4.4

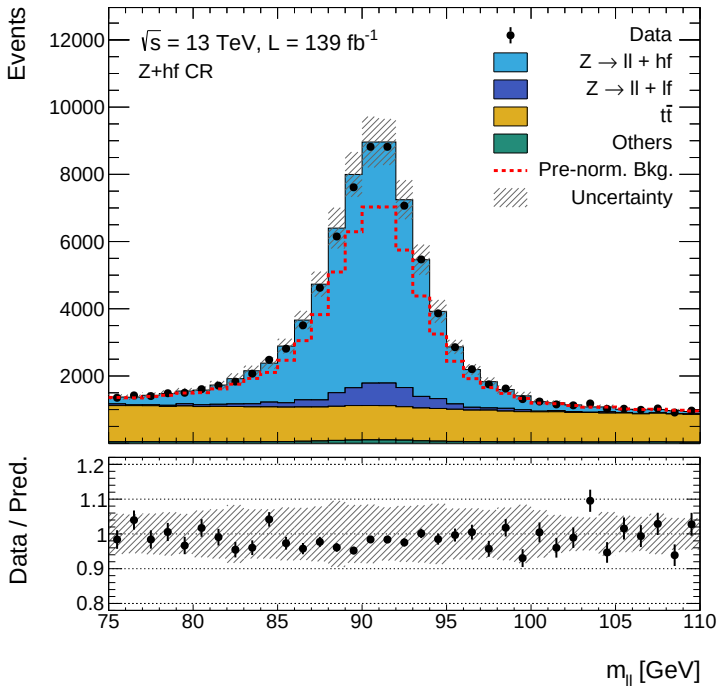


Figure 4.7: Data–background comparison of the $m_{\ell\ell}$ distribution with the normalisation factor of $Z+hf$ (1.4) and $t\bar{t}$ (0.96) applied. The uncertainty includes statistical and experimental uncertainties of the simulated samples. Distribution of the total background before applying the normalisation factors ('Pre-norm. Bkg.') is displayed in the dashed red line.

- Tighter $m_{\ell\ell}$ selection and additional $\Delta(\ell,\ell)$ selection to further target at $Z+hf$ events: $85 < m_{\ell\ell} < 97$ GeV, $\Delta R(\ell,\ell) < 1.5$;
- Select m_{bb} mass window to get close to signal region topology: $50 < m_{bb} < 150$ GeV;
- Require small $E_{\mathrm{T}}^{\mathrm{miss}}$ to reject contamination of $HH \to bbZZ$ events: $E_{\mathrm{T}}^{\mathrm{miss}} < 45$ GeV.

Figure 4.8 shows the distribution of BDT and some exemplary PNN scores, where the $Z+hf$ is found to be well modelled.

The estimations of the generic Z+jets and $Z+hf$ are validated in the validation regions listed in Table B.1 [9]. Data–background comparison of the MVA input variables in these VRs is shown in Figure B.3 and Figure B.4.

4.3.2 Top-quark pair

A pair of top-quark can produce two real τ_{had}'s via a $t\bar{t} \to W^+bW^-\bar{b}(W^+ \to \tau^+\nu, W^- \to \tau^-\bar{\nu})$ decay chain. This process is named as the true-τ_{had} $t\bar{t}$ process [10]. If one or both of the W bosons decay hadronically, one or both of the reconstructed τ_{had} candidates are likely to be misidentified from jets initiated from $W \to q\bar{q}'$. These processes are called the fake-τ_{had} $t\bar{t}$ process. We use *Fake*

[9] The validations are collected in Appendix B.3, so as for the other background processes described below.

[10] If both W bosons decay to real electrons or muons. It's also considered as true-τ_{had} $t\bar{t}$, but the fraction of this process is small because of the electron veto (see Section 3.2)

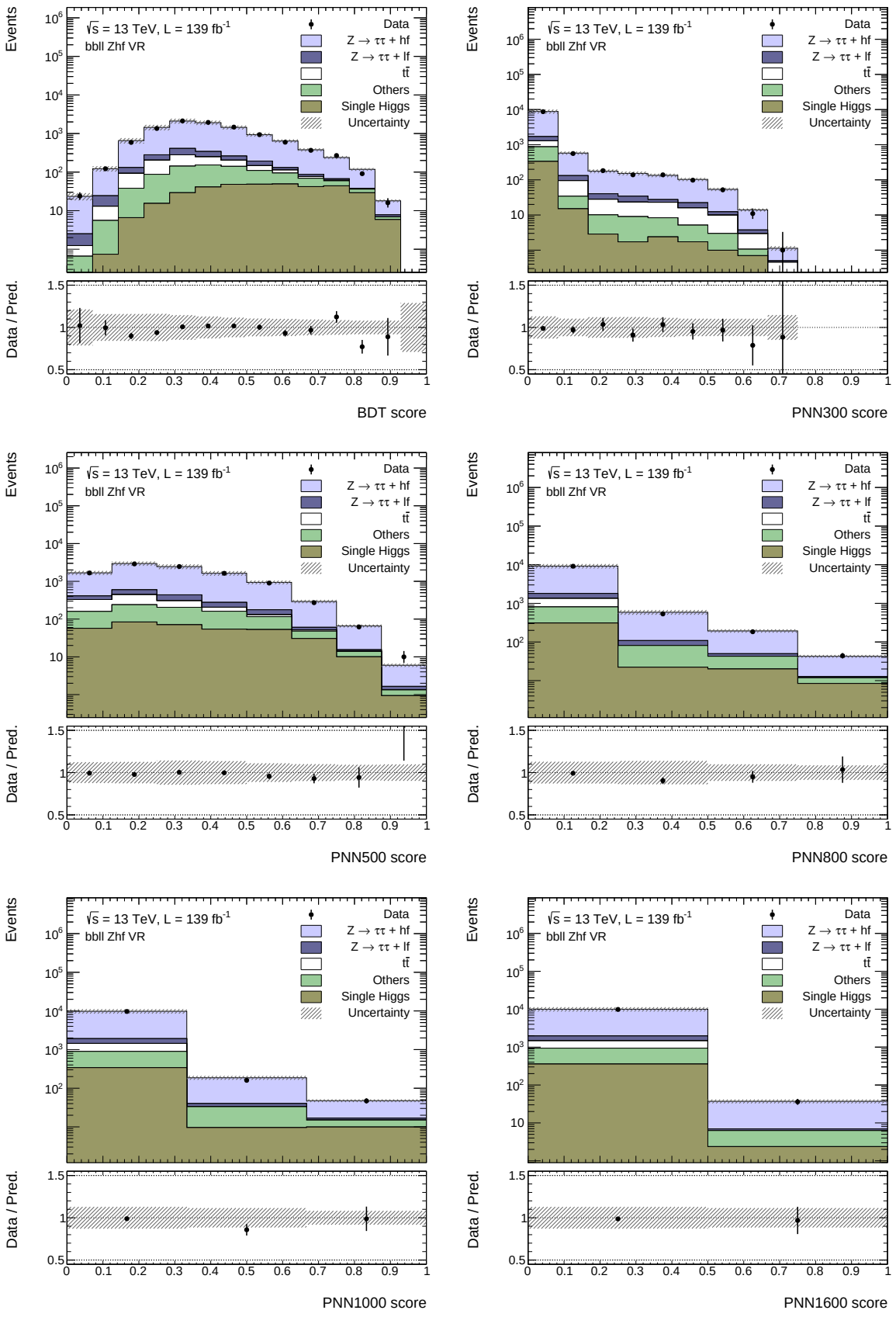


Figure 4.8: Distribution of BDT and several exemplary PNN scores in the $b\bar{b}\ell^+\ell^-$ $Z+hf$ VR. The $Z+hf$ is scaled by a normalisation factor of 1.2 so that the total background is compatible with the data. The binnings are the same as those used in the SR. In the lower panel, the ratio of data and the total background prediction is displayed in black dots. The hatched area indicates the statistical uncertainty of the simulated samples and the $Z+hf$ theoretical uncertainties are described in Section 4.4.2.2.

Rates (FRs) to quantify the rate for jets to be misidentified as τ_{had}.

4.3.2.1 Reweighting of the simulated $t\bar{t}$ samples

The POWHEG BOX V2 generates $t\bar{t}$ at NLO precision. According to the $t\bar{t}$ cross section measurements [146, 147], the missing high order calculation leads to discrepancies between the predicted and observed p_{T} distributions of the top-quark and its decay products. The discrepancy is especially pronounced in the high jet multiplicity and high p_{T} region. This can cause problems when a significant amount of $t\bar{t}$ is subtracted in the derivation or application of the fake-τ_{had} estimations, for both $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channel. Thus, $t\bar{t}$ needs to be corrected in prior to fake-τ_{had} estimations.

The discrepancy can be mitigated by the NNLO QCD corrections to generator-level top-quark p_{T} and $t\bar{t}$ invariant mass distributions [148]. Instead of using the NNLO corrections, we perform an in situ $t\bar{t}$ correction using events in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. The $\tau_{\text{lep}}\tau_{\text{had}}$ channel is used because it is difficult to define a $t\bar{t}$-enriched region in $\tau_{\text{had}}\tau_{\text{had}}$ channel. The signal region selections of the $\tau_{\text{lep}}\tau_{\text{had}}$ channel are described in Section 4.7

A control region called $t\bar{t}$-noID-CR is defined to derive the $t\bar{t}$ correction reweighting factors (named as $t\bar{t}$ reweighting). The selections of $t\bar{t}$-noID-CR closely follow the single-lepton-trigger category of the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, but without requiring τ_{had}-ID of the τ_{had} candidate. This is because the $t\bar{t}$ reweighting is designed to correct both true and fake-τ_{had} $t\bar{t}$ events, thus it must be independent of τ_{had}-ID. Besides, the $t\bar{t}$-noID-CR requires $m_{bb} > 150$ GeV and transverse mass of W boson [11] $m_{\text{T}}^{W} > 40$ GeV. These two requirements are deployed to select a $t\bar{t}$-enriched region and to eliminate the contamination from potential multi-jet process, respectively.

Firstly, the $t\bar{t}$ is normalised such that the overall background is compatible with the data, the normalisation factor is 0.93 ± 0.023. The uncertainty is obtained by the statistical uncertainty of data and simulated samples (0.1%), the subtraction of non-$t\bar{t}$ processes [12] (1.4%) and potential missing multi-jet in the subtraction (2%) [146].

Next, a two-dimensional shape correction parameterized by the jet multiplicity (N_{jets}) and the scalar sum of the transverse momentum of all visible final state objects ($H_{\text{T}} = \sum p_{\text{T}}^{\text{jets}} + p_{\text{T}}^{e/\mu} + p_{\text{T}}^{\tau_{\text{had}}}$) is derived by

$$f_{t\bar{t}}(H_{\text{T}}, N_{\text{jets}}) = \frac{N(\text{data}) - N(\text{MC, non-}t\bar{t})}{N(\text{MC}, t\bar{t})} \tag{4.3}$$

where N stands for the number of events. The reweighting factors are visualised in Figure 4.11. The following sources of systematic uncertainties are considered for the reweighting:

- Full difference between reweighted and not reweighted $t\bar{t}$ samples;
- Statistical uncertainties of data and simulated samples;

[11] $m_{\text{T}}^{W} = \sqrt{2 p_{\text{T}}^{\ell} E_{\text{T}}^{\text{miss}} (1 - \cos\Delta\phi(p_{\text{T}}^{\ell}, E_{\text{T}}^{\text{miss}}))}$, where ℓ represents an electron or a muon.

[12] A relative uncertainty of 20% on single top and 40% on all of the other minor backgrounds (mostly W+jets).

- Residual discrepancy between data and background in the τ_{had} p_{T} distribution (non-closure uncertainty).

The first term is the dominant uncertainty. The third term is introduced given the fact that the reweighting parameterized with H_{T} is not optimal for τ_{had} p_{T} because τ_{had} p_{T} contributes less to the H_{T}. It is derived as a function of τ_{had} p_{T} and n-prong by applying the reweighting back to the $t\bar{t}$-noID-CR (so-called closure check), as shown in Figure 4.10. The background modelling of $t\bar{t}$ after the reweighting is compared with that before the reweighting in Figure 4.9 for the distributions of variables that are sensitive to the reweighting, in the $t\bar{t}$-noID-CR. Same reweighting is applied on true-τ_{had} and fake-τ_{had} $t\bar{t}$ simulated events. The reweighting shows a good closure in $t\bar{t}$-noID-CR.

The $t\bar{t}$ reweighting is validated in two VRs, one of them checks the independence of the reweighting on m_{bb}, and the other one checks exclusively the performance of the reweighting on true-τ_{had} $t\bar{t}$ events. The definition of those VRs can be found in Table B.2, where different selections with respect to $t\bar{t}$-noID-CR are explained. Data–background comparison of the distributions of variables that are sensitive to $t\bar{t}$ reweighting are depicted in Figure B.5–B.6.

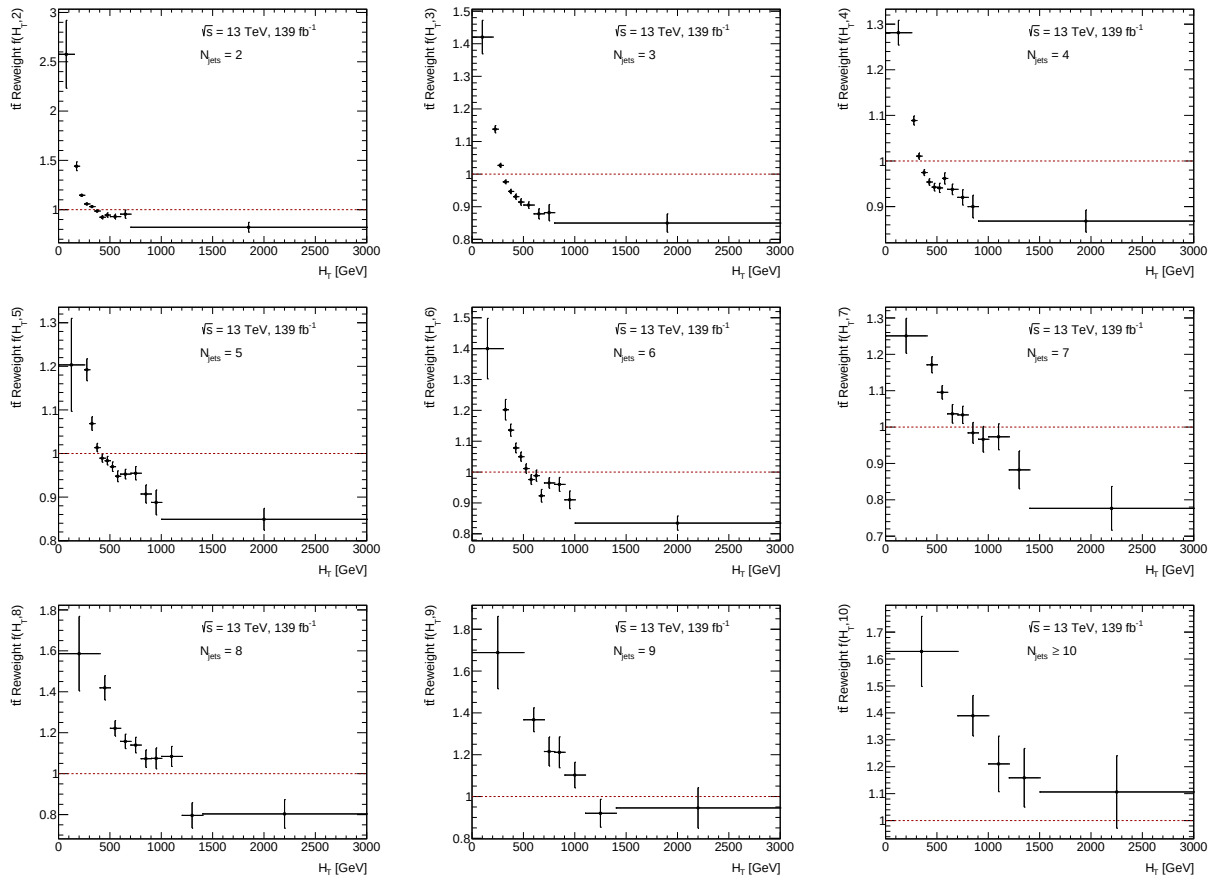
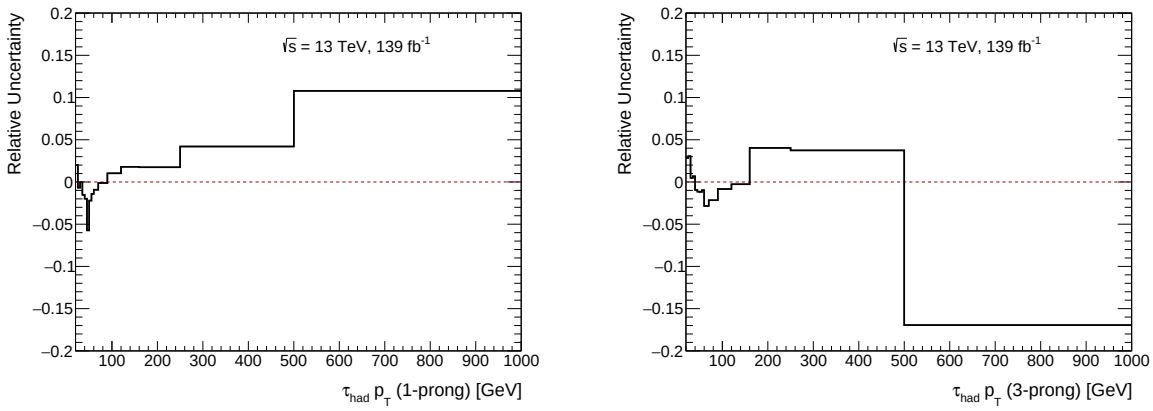


Figure 4.9: The $t\bar{t}$ reweighting factors as functions of H_{T} and N_{jets}. The uncertainties shown are the statistical uncertainties of data and simulated samples.



(a) 1-prong (b) 3-prong

Figure 4.10: Non-closure uncertainties of the $t\bar{t}$ reweighting as functions of τ_{had} p_{T}, derived separately for 1-prong and 3-prong τ_{had}.

4.3.2.2 Estimation of fake-τ_{had} $t\bar{t}$ background

The fake rate of the fake-τ_{had}'s in $t\bar{t}$ events can be measured after applying the $t\bar{t}$ reweighting on both the true- and fake-τ_{had} $t\bar{t}$ events. There are two approaches to measure it — MC-based approach and data-driven approach. In the MC-based approach, the fake rate is measured by the simulated fake-τ_{had} $t\bar{t}$ events (fake-τ_{had} $t\bar{t}$ MC). The *FR* is defined as

$$FR_{\text{MC}} = \frac{N_{\text{pass}}(\text{fake-}\tau_{\text{had}}\ t\bar{t}\ \text{MC})}{N_{\text{total}}(\text{fake-}\tau_{\text{had}}\ t\bar{t}\ \text{MC})}, \tag{4.4}$$

where N_{pass} stands for the number of events that pass the *Loose* τ_{had}-ID selection, and N_{total} is the total number of events before applying the τ_{had}-ID selection. Note that there is one τ_{had} candidate per event in the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. In the data-driven approach, the *FR* is defined as

$$FR_{\text{data-driven}} = \frac{N_{\text{pass}}(\text{data}) - N_{\text{pass}}(\text{other MC})}{N_{\text{total}}(\text{data}) - N_{\text{total}}(\text{other MC})}, \tag{4.5}$$

where 'other MC' stands for MC events of all other processes except fake-τ_{had} $t\bar{t}$. After applying the $t\bar{t}$ reweighting, the denominator of Equation (4.4) and Equation (4.5) are equal. Therefore, a scale factor called the fake-τ_{had} $t\bar{t}$ *SF* (labelled as *SF* for clarity in the following) can be defined in the form of

$$SF = \frac{FR_{\text{data-driven}}}{FR_{\text{MC}}} = \frac{N_{\text{pass}}(\text{data}) - N_{\text{pass}}(\text{other MC})}{N_{\text{pass}}(\text{fake-}\tau_{\text{had}}\ t\bar{t}\ \text{MC})}. \tag{4.6}$$

The *SF* can be used to correct the simulated fake-τ_{had} $t\bar{t}$ samples in the $\tau_{\text{had}}\tau_{\text{had}}$ channel. As indicated in Equation (4.6), unlike the measurement of the fake rate, the measurement of *SF* only require events that pass τ_{had}-ID. A control region labelled as $t\bar{t}$-ID-CR can be defined by applying the *Loose* τ_{had}-ID selection on top of the $t\bar{t}$-noID-CR. Figure 4.12 illustrates the data-MC background comparison in $t\bar{t}$-ID-CR.

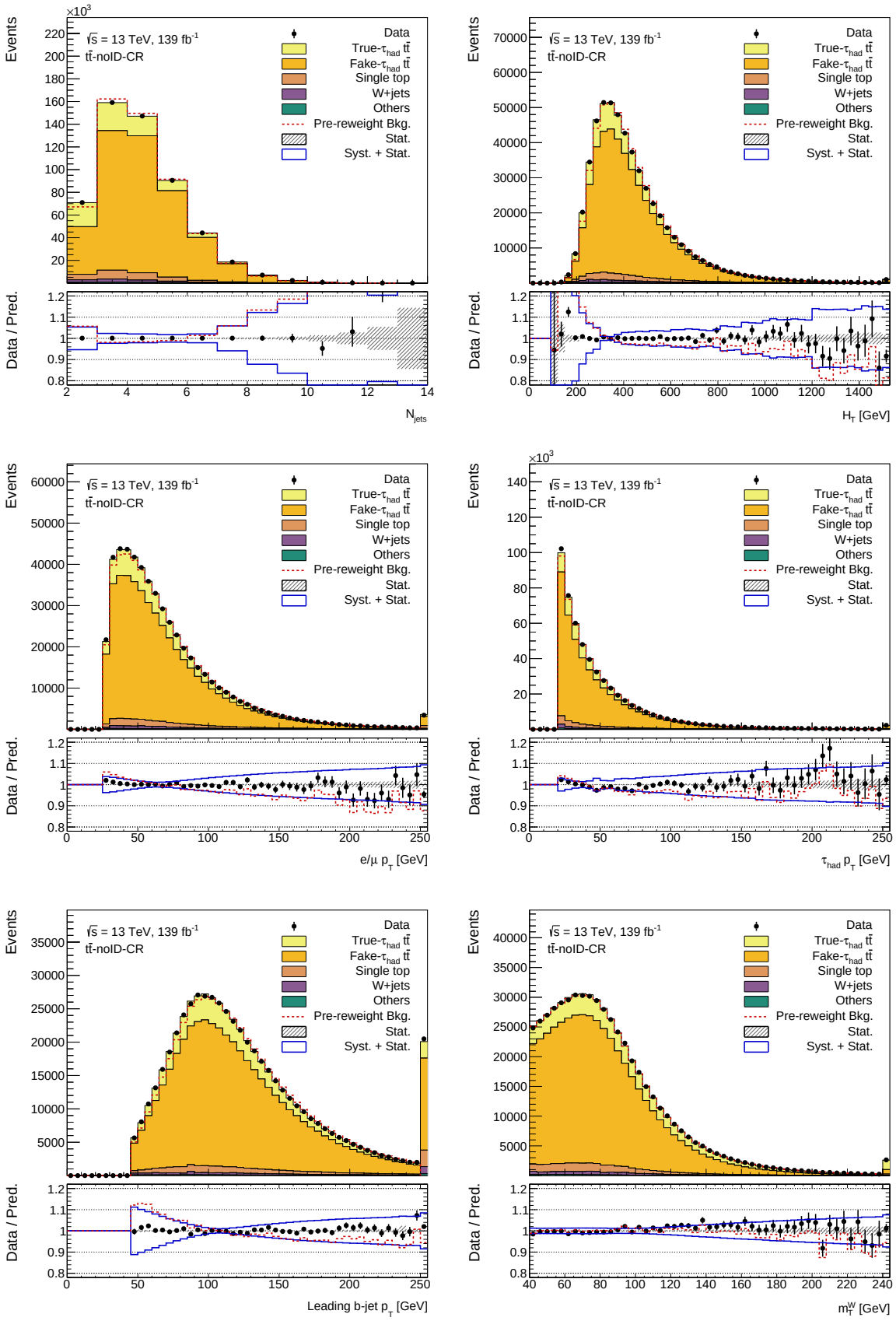


Figure 4.11: Comparison of the distributions of the kinematics of final state objects before and after $t\bar{t}$ reweighting. The shape of total background before reweighting (Pre-reweight) is displayed in dashed red line. The statistical uncertainty of simulated samples is shown in hatched boxes (Stat.), while the impact of systematic uncertainties of $t\bar{t}$ reweighting is indicated in the blue lines (Syst. + Stat.). Overflow of events is included in the last bin.

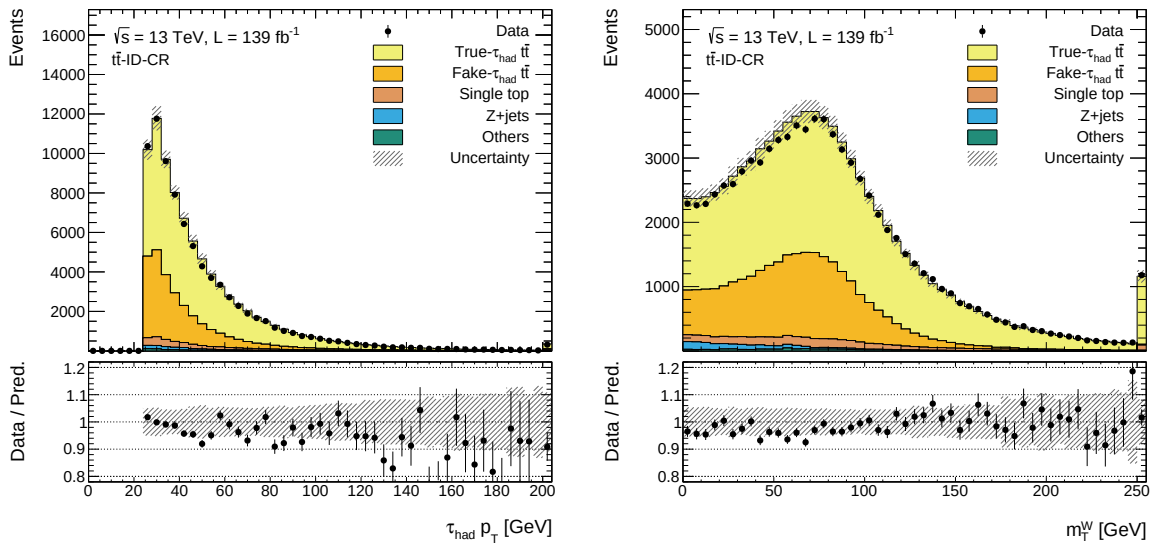


Figure 4.12: Data-MC background comparison of τ_{had} p_{T} (left) and m_{T}^{W} (right) distributions in $t\bar{t}$-ID-CR. $t\bar{t}$ reweighting is applied to both true and fake-τ_{had} $t\bar{t}$ processes. The uncertainty includes statistical and experimental uncertainties of the simulated samples.

The *SF* needs to be measured for τ_{had}'s with different trigger selections according to the τ_{had}'s that are used in the $\tau_{\text{had}}\tau_{\text{had}}$ channel. For the events that pass a specific τ_{had} trigger, the offline τ_{had} must be trigger-matched. The *SF*s are only measured for τ_{had}'s that pass triggers with 25 GeV p_{T} threshold. For τ_{had} that pass triggers with higher p_{T} thresholds, the corresponding *SF* as 25 GeV threshold triggers are applied. The generator-level truth origins of the τ_{had} candidates with different kinematics and n-prongs can be different. Thus, the *SF*s are also measured in categories of τ_{had} p_{T} and n-prong [13]. The different triggers and the corresponding *SF*s are summarised in Table 4.7.

Table 4.7: Different triggers considered in the *SF* measurement. The available periods of the trigger selections and abbreviations of the *SF*s are listed. Each *SF* is a function of τ_{had} p_{T} and n-prong. Note that there is a logical *OR* between the `medium1_tracktwoEF` and `mediumRNN_tracktwoMVA` triggers starting from period K of 2018 data-taking.

Trigger	Available period	Abbreviations	τ_{had} to be applied in $\tau_{\text{had}}\tau_{\text{had}}$
No trigger requirement	All periods	*SF*(off)	sub-leading-p_{T} τ_{had} in STT events
`HLT_tau25_medium1_tracktwo`	All periods	*SF*(tau25)	leading-p_{T} τ_{had} in STT events and all τ_{had} in DTT events
`HLT_tau25_medium1_tracktwoEF`	2018	*SF*(tau25 EF)	same as above
`HLT_tau25_medium1_tracktwoEF` *OR* `HLT_tau25_mediumRNN_tracktwoMVA`	$\geq$ 2018 period K	*SF*(tau25 EForMVA)	same as above

The *SF*s are measured separately for each trigger-selection scenario by the likelihood fits, using the data in the corresponding available period. In each fit, and for each p_{T}, n-prong category, one *SF* is assigned as a normalisation factor to the fake-τ_{had} $t\bar{t}$ events. The fit is performed simultaneously across all p_{T} and n-prong categories using the m_{T}^{W} distribution. The m_{T}^{W} is used because

[13] The τ_{had} p_{T} are categorised in [25, 30), [30, 35), [35, 40), [40, 45), [45, 55), [55, 70), [70, ∞) for 1-prong τ_{had}, and [25, 30), [30, 40), [40, 50), [50, 70), [70, ∞) for 3-prong. The unit is GeV.

the factions of true- and fake-τ_{had} $t\bar{t}$ varies significantly from low to high m_{T}^{W} values, as shown in Figure 4.13 in the case of measuring the SF(off) in the categories of $p_{\text{T}} \in [40, 45)$ GeV and 1- or 3-prong. The distributions in other p_{T} categories are similar. The experimental uncertainties, the $t\bar{t}$ reweighting uncertainties and the theoretical uncertainties in the other minor processes are profiled. The normalisation of $t\bar{t}$ is a free parameter. Figure 4.14 shows the best-fit values and 1σ confidence intervals of the $t\bar{t}$ SFs of τ_{had} candidates with different trigger selections, as functions of τ_{had} p_{T} and n-prong. Good agreement is found between data and post-fit background, as shown in Figure 4.13, for $t\bar{t}$-ID-CR with different trigger selections.

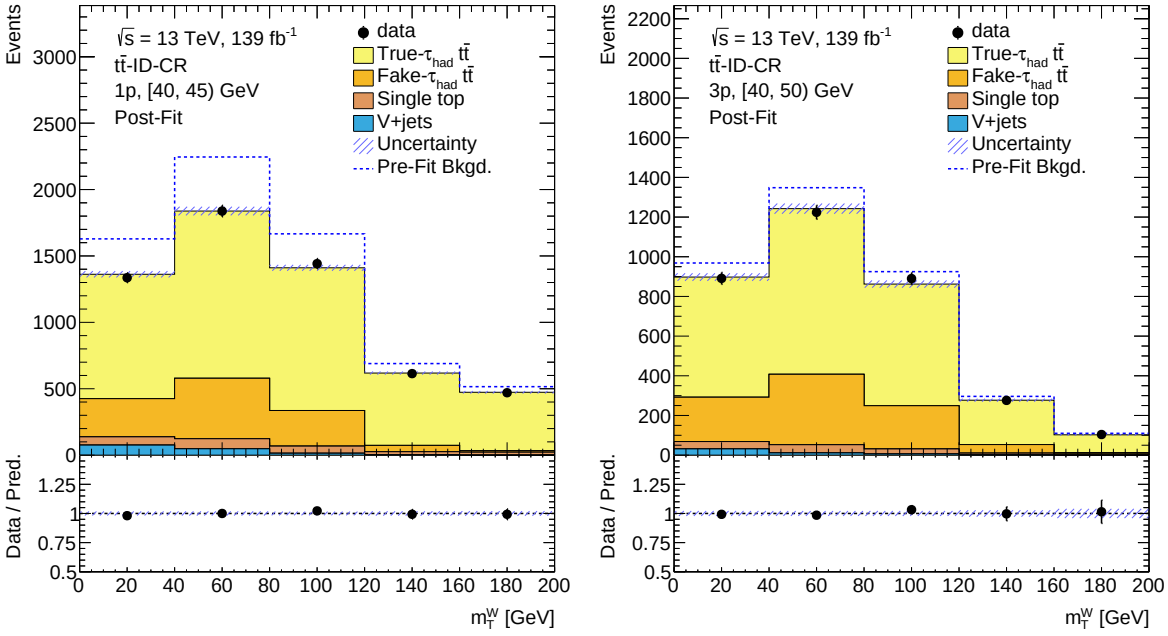


Figure 4.13: The m_{T}^{W} template used to measure SF(off) in the categories of τ_{had} $p_{\text{T}} \in [40, 45)$ and 1- (left) or 3-prong (right). The post-fit background distribution is shown in stacked histogram while the pre-fit background is shown in the dashed blue line. The uncertainty band includes all sources of uncertainties as mentioned in the text.

The SFs in each fit are not independent because the (p_{T}, n-prong) categories are correlated by the nuisance parameters related to $t\bar{t}$ normalisation and systematic uncertainties. To preserve the correlation of SFs in the application to the $\tau_{\text{had}}\tau_{\text{had}}$ channel, eigen-variations are calculated from the covariance matrix of the SFs. The two largest eigen-variations [14] on SF(tau25) for 1-prong and 3-prong τ_{had} are illustrated in Figure 4.15. Additionally, a relatively small uncertainty accounting for difference between SFs derived by triggers with a p_{T} threshold of 25 GeV and 35 GeV is considered.

The SFs are also measured for anti-τ_{had} offline selection. The procedure is very similar as the case of passing *Loose* τ_{had}-ID selection. These SFs are applied to correct the fake-τ_{had} $t\bar{t}$ distribution in the anti-ID region of multi-jet FF application (see Section 4.3.3).

The estimations of true and fake-τ_{had} $t\bar{t}$ are validated in the VRs listed in Table B.3 with the

[14] They correspond to the two largest eigenvalues of the covariance matrix.

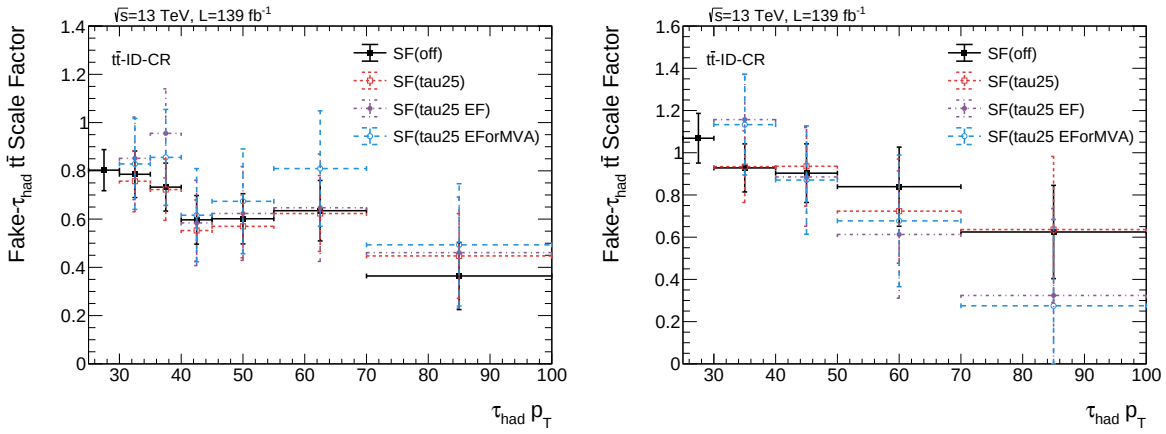


Figure 4.14: The best-fit values and 1σ confidence intervals of the $t\bar{t}$ *SF*s of τ_{had} candidates with different trigger selections, as functions of τ_{had} p_{T} and n-prong — 1-prong (left) and 3-prong (right). The error bars indicate the total uncertainty of the scale factors.

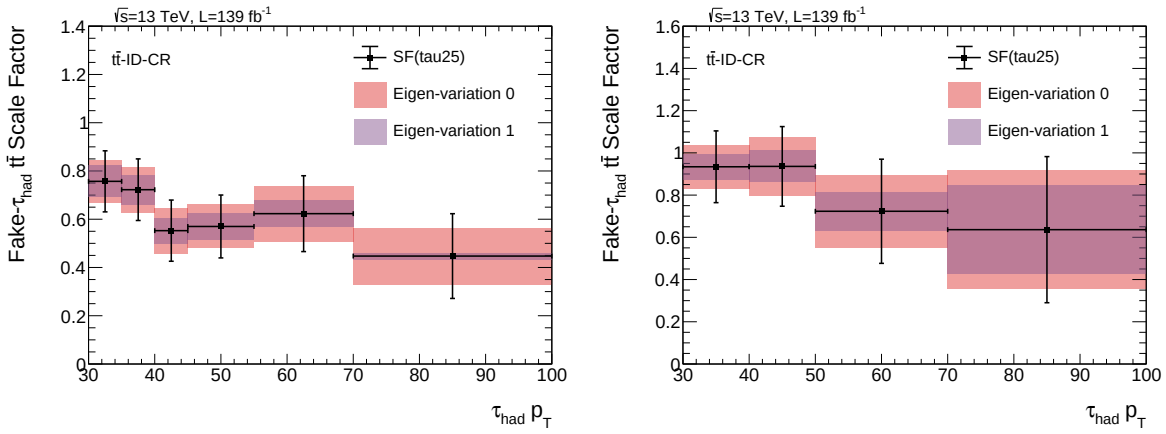


Figure 4.15: The two largest eigen-variations on *SF*(tau25) for 1- (left) and 3-prong (right) τ_{had}, which are shown in red and blue error bands. The black vertical error bar indicates the total uncertainty of *SF*(tau25) determined in the likelihood fits.

different selections compared with SR listed. Note that in all the $\tau_{\text{had}}\tau_{\text{had}}$ channel VRs and SR, the $t\bar{t}$ reweighting is not applied to be consistent with $\tau_{\text{lep}}\tau_{\text{had}}$ and $b\bar{b}\ell^+\ell^-$ final states for later combination. Alternatively, the theoretical uncertainty is employed to add degrees of freedom to manipulate the shape of $t\bar{t}$. It is also found that the $t\bar{t}$ reweighting does not have obvious shape effects on the $\tau_{\text{had}}\tau_{\text{had}}$ channel. Data–background comparison of the MVA input variables in the $t\bar{t}$ VRs are depicted in Figure B.7.

4.3.3 Multi-jet

Multi-jet events can be selected as signal candidates when two quark/gluon-initiated jets are misidentified as τ_{had}'s. This process is estimated by a data-driven Fake Factor (*FF*) method. The *FF* is the ratio of multi-jet events in ID and anti-ID region. The regions with both τ_{had} passing the *Loose* τ_{had}-ID are called ID regions, while the regions with one of the τ_{had} being an anti-τ_{had} are called anti-ID regions. The anti-τ_{had} objects are also required to be trigger-matched. If multiple anti-τ_{had} objects are found, one of them is selected randomly. The ID and anti-ID region that are used to calculate the *FF*s are defined based on the SR but with a 1-*b*-tagged same-sign (SS) selection. Here, the SS represents the same-sign of the electric charges of the two τ_{had} candidates. The *FF*s are measurement by subtracting the simulated non-multi-jet events from data

$$FF = \frac{N_{\text{ID}}(\text{data}) - N_{\text{ID}}(\text{non-multi-jet})}{N_{\text{Anti-ID}}(\text{data}) - N_{\text{Anti-ID}}(\text{non-multi-jet})}, \tag{4.7}$$

where N denotes for the number of events and the subscript stands for the regions. Figure 4.16 shows the distributions of τ_{had} p_{T} for 1-prong τ_{had}'s in 1-*b*-tagged SS ID and anti-ID regions. The discrepancy between data and MC prediction is supposed to be the missing multi-jet process. For 3-prong τ_{had}'s, the distributions are similar.

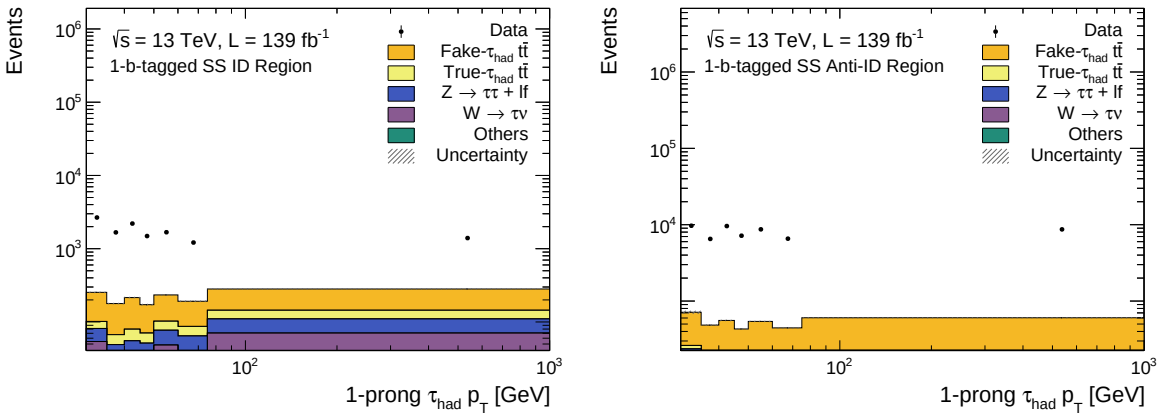


Figure 4.16: Distributions of τ_{had} p_{T} for 1-prong τ_{had}'s in 1-*b*-tagged SS ID and anti-ID regions. The discrepancy between data and MC prediction indicates the missing simulation of multi-jet process which will be estimated by the fake factor method.

The *FF*s are derived as functions of τ_{had} p_{T}, η, n-prong and triggers (`STT` and `DTT`). They are

also split by years to account for the different trigger configurations between data-taking periods. For the DTT category, the *FF*s are originally derived separately for p_{T}-leading and sub-leading τ_{had} candidates, but the *FF*s are found to be compatible so the average of them is used. Different *FF*(s) are derived in the barrel ($0 < |\eta| < 1.37$) and endcap ($1.52 < |\eta| < 2.5$) areas split by η according to the geometry of the EM calorimeter. Due to the limited number of events of STT category, the *FF*s are calculated inclusively in τ_{had} p_{T} and η but separately based on whether the τ_{had} candidate is p_{T} leading or sub-leading. These *FF*s are summarised in Figure 4.17.

The *FF*s are derived in 1-*b*-tagged SS regions and applied to the 2-*b*-tagged SS and OS regions. A set of Transfer Factors (*TF*s) is used to account for the extrapolation from 1- to 2-*b*-tagged SS region. It is defined as the ratio of *FF*s measured in 2-*b*-tagged and 1-*b*-tagged SS regions

$$TF = \frac{FF_{2-b\text{-tagged}}}{FF_{1-b\text{-tagged}}}. \tag{4.8}$$

Other than the *FF*s, *TF*s are measured inclusively in trigger categories and τ_{had} p_{T}, η but separately based on whether the τ_{had} candidate is p_{T} leading or sub-leading. The *TF*s are shown in Figure 4.18. Finally, the estimation of the shape and normalisation of multi-jet background in the SR (2-*b*-tagged OS ID region) is constructed by applying the *FF*s and *TF*s to the events in 2-*b*-tagged OS anti-ID region, i.e.,

$$\begin{aligned} N_{\mathrm{SR}}(\text{multi-jet}) = {} & TF \\ & \times FF_{1-b\text{-tagged}} \\ & \times \left[N_{\text{Anti-ID region}}(\text{data}) - N_{\text{Anti-ID region}}(\text{non-multi-jet})\right], \end{aligned} \tag{4.9}$$

where 'Anti-ID region' stands for 2-*b*-tagged OS anti-ID region.

Figure 4.19 shows the distribution of τ_{had} p_{T} for 1-prong and 3-prong τ_{had}'s in 2-*b*-tagged OS anti-ID region. A significant amount of fake-τ_{had} $t\bar{t}$ is subtracted in this region, which is corrected by dedicated anti-ID fake-τ_{had} $t\bar{t}$ *SF*s mentioned in Section 4.3.2.2. Besides, the extrapolation from SS to OS region is considered by introducing an uncertainty based on the comparison of *FF*s calculated in 1-*b*-tagged SS and 1-*b*-tagged OS [15] regions. Other sources of uncertainty are from the statistical uncertainties in the calculation of *FF*s and *TF*s and from the subtracted non-multi-jet samples.

Good closure is found when the *FF*s are applied back to the 1-*b*-tagged SS regions, as shown in Figure 4.20. The estimation of multi-jet is validated in the regions defined in Table B.4. Distributions of MVA input variables in those regions are shown in Figure B.8 and Figure B.9.

[15] With additional requirements of $m_{\tau\tau}^{\mathrm{MMC}} > 110$ GeV and $E_{\mathrm{T}}^{\mathrm{miss}}$ significance < 3 to selected high multi-jet purity region, see Appendix B.3.

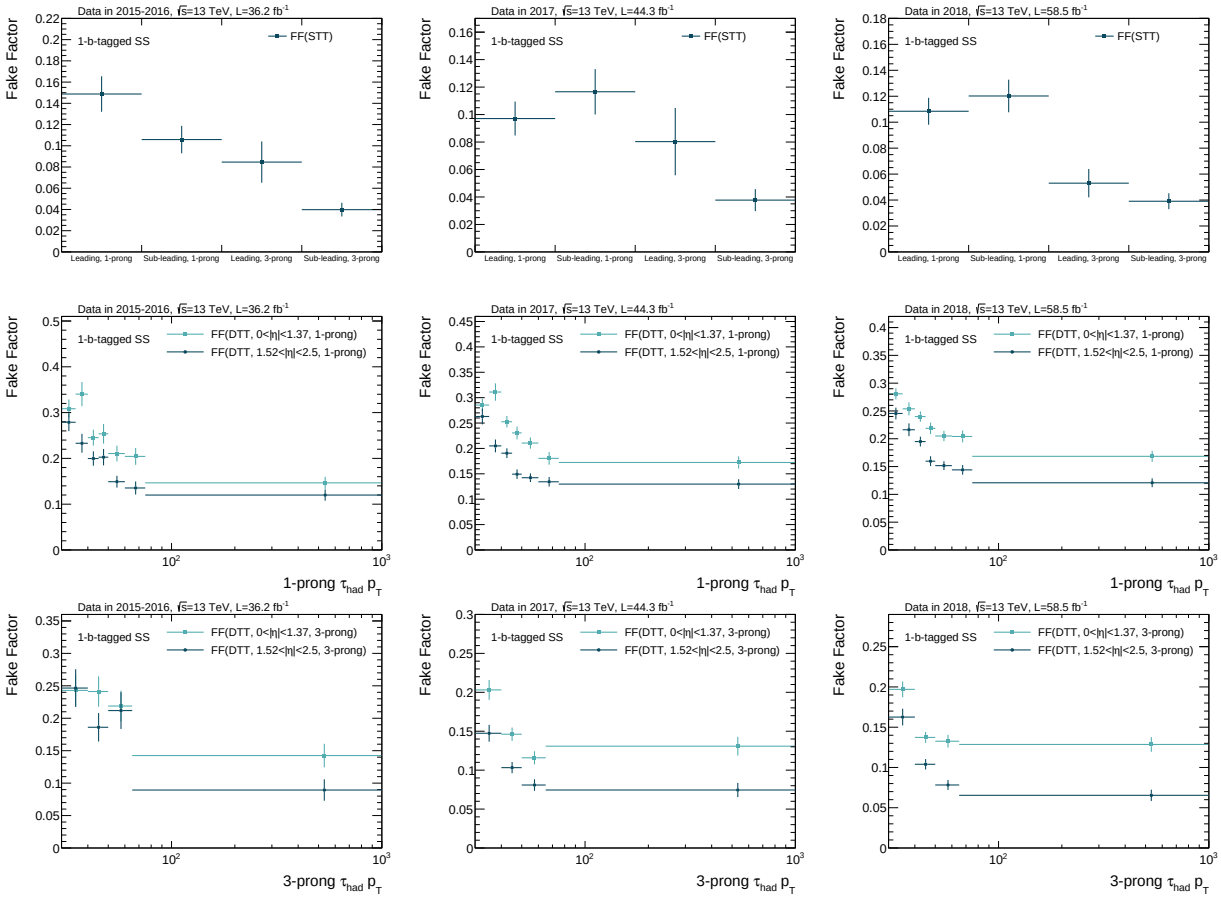


Figure 4.17: Multi-jet fake factors for STT (first row) and DTT (other rows) categories and different data-taking years (shown in the label). The parameterization of the fake factors is described in the text.

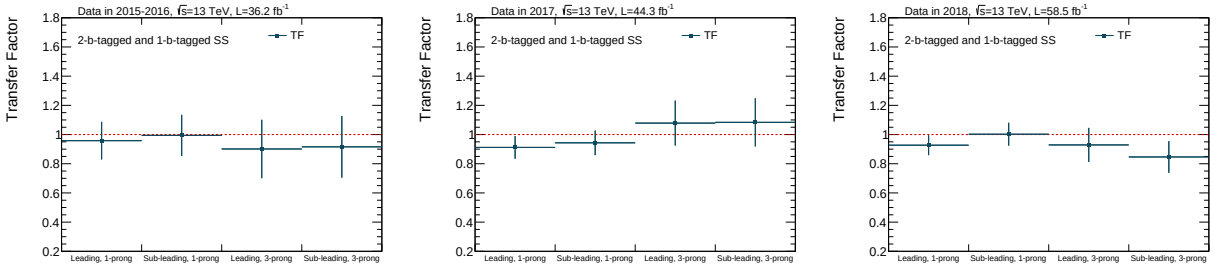


Figure 4.18: The 1-*b*-tagged to 2-*b*-tagged fake-factor transfer factors in different years (shown in the label). The parameterization of the fake factors is described in the text.

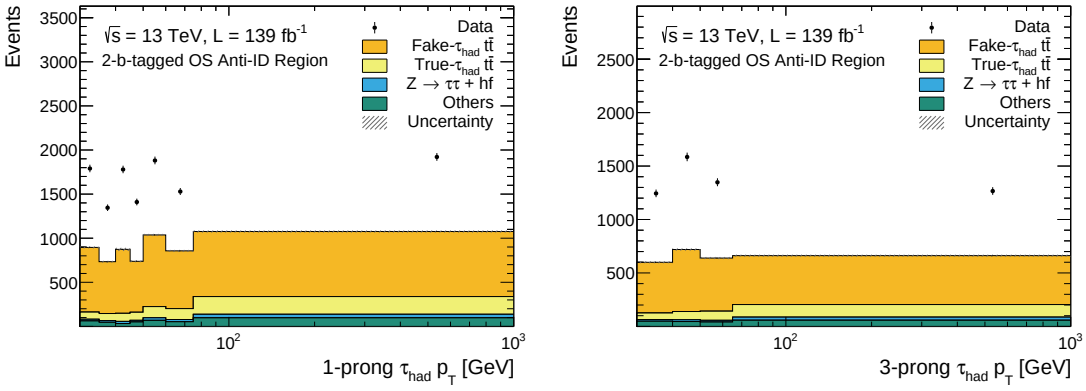


Figure 4.19: Distributions of τ_{had} p_{T} for 1- and 3-prong τ_{had}'s in 2-b-tagged OS anti-ID region. The difference between data and MC simulation is used as template to estimate the multi-jet contribution in the SR.

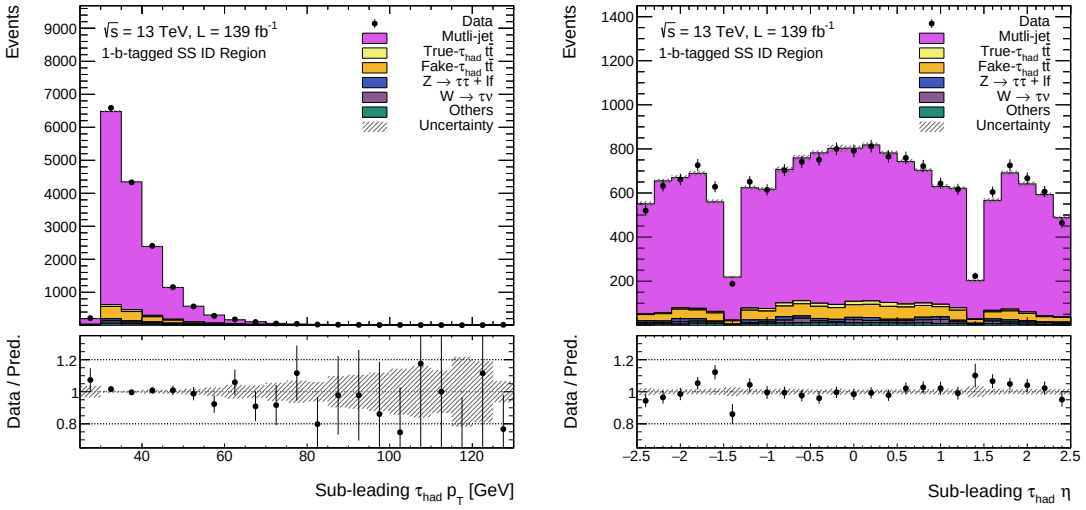


Figure 4.20: Distributions of τ_{had} p_{T} and η in the 1-b-tagged SS region to check the closure of the *FF*s. Sub-leading refer to the ordering with respect to p_{T}. The uncertainty includes the statistical uncertainty of the simulated samples in the ID-region and statistical uncertainty of the multi-jet template in the anti-ID region, as well as the experimental uncertainties.

4.4 Systematic uncertainties

4.4.1 Overview of uncertainties

Uncertainty originates from several sources. A large fraction is from the statistical nature of the collision data. All the other uncertainties are considered as systematic uncertainties. These uncertainties can affect the total event yields in the SR and can also distort the shape of the distributions of, e.g., the MVA scores that are used in the binned likelihood fit to determine the statistical results, where the binning is decided based on the algorithm discussed in Section 4.2.4.4. The contribution of each source of uncertainties to the total uncertainty of the parameter of interest is presented in Section 4.6.

Firstly, the MC simulation samples are generated with finite number of events, therefore the statistical uncertainty needs to be taken into account.

Besides, the simulated samples are also generated with fixed precision of calculations and specific configurations of the generators. As described in Section 2.4, the total and differential cross sections depend on those setting and uncertainties need to be considered for each components of the simulation. For ATLAS physics analyses, it is recommended to tune the parameters within certain range or compare the outcomes of different generators to determine the impact of the systematic uncertainties from the theoretical aspects. For a given observable N, the uncertainty is defined by the difference between the alternative sample or configuration and the nominal one,

$$\Delta N = N_{\text{alternative}} - N_{\text{nominal}}. \tag{4.10}$$

The uncertainties in total cross sections that are used to normalise the simulated samples are considered [16], as well as the uncertainties in the analysis acceptance times selection efficiency $\mathscr{A} \times \epsilon$. The latter can affect both total (normalisation effect) and differential yield (shape effect) of the simulated events. The normalisations of $Z + hf$ and $t\bar{t}$ are determined in the likelihood fit and are driven by the $Z + hf$ CR. Following the procedure described in Ref. [149], relative normalisation uncertainties on $\mathscr{A} \times \epsilon$ is derived to account for the residual differences in the extrapolation of normalisation from CR to SR (CR $\rightarrow$ SR), which is defined as

$$\Delta\mathscr{A} \times \epsilon_{\text{CR}\rightarrow\text{SR}} = \left(\frac{\mathscr{A} \times \epsilon_{\text{SR}}}{\mathscr{A} \times \epsilon_{\text{CR}}}\right)_{\text{alternative}} - \left(\frac{\mathscr{A} \times \epsilon_{\text{SR}}}{\mathscr{A} \times \epsilon_{\text{CR}}}\right)_{\text{nominal}}. \tag{4.11}$$

The remaining shape effects are derived separately and they vary coherently with the normalisation effects. The uncertainties for individual processes are elaborated in Section 4.4.2.

The uncertainties of background estimation are discussed in Section 4.3. For $Z + hf$ and $t\bar{t}$ the uncertainties of the normalisation factors are given by the likelihood fit. The uncertainties in fake-τ_{had} $t\bar{t}$ and multi-jet estimations are summarised in Table 4.8. The sources of uncertainties are

[16] Not for $Z + hf$ and $t\bar{t}$ as their normalisations are determined by the likelihood fit.

reviewed in the descriptions.

Table 4.8: Summary of systematic uncertainties related to fake-τ_{had} $t\bar{t}$ and multi-jet estimations.

Name of uncertainty	Description
FAKETTBAR_OFFL_EIGEN[0-11] FAKETTBAR_TAU25_EIGEN[0-9] FAKETTBAR_TAU25EF_EIGEN[0-9] FAKETTBAR_TAU25RNN_EIGEN[0-9]	Uncertainties in fake-τ_{had} $t\bar{t}$ estimated by the *SF*s derived with different 'trigger+offline' conditions of τ_{had}. The number indicates multiple eigen-variations from the diagnonalisation of the covariant mastrx of *SF*s.
FAKETTBAR_TAU25_TAU35_DIFF FAKETTBAR_TAU25EF_TAU35EF_DIFF FAKETTBAR_TAU25RNN_TAU35RNN_DIFF	Uncertainties that account for the difference of *SF*s calculated with τ_{had} with p_{T} thresholds of 25 GeV and 35 GeV.
FF_Stat_HadHadSR	Statistical uncertainty of the *FF*s
FF_OSSS	SS to OS region extrapolation uncertainty
FF_TRUE_TTBAR_SUBTRACTION	Uncertainty of the subtracted true-τ_{had} $t\bar{t}$
FAKETTBAR_ANTITAU_DIFF FAKETTBAR_ANTITAU_OFFL_EIGEN0 FAKETTBAR_ANTITAU_TAU25_EIGEN0 FAKETTBAR_ANTITAU_TAU25EF_EIGEN0 FAKETTBAR_ANTITAU_TAU25RNN_EIGEN0	Uncertainties in the subtracted fake-τ_{had} $t\bar{t}$ estimated by the *SF*s derived with different 'trigger+offline' conditions of the anti-τ_{had}. The full difference between applying and not applying the *SF*s is considered as uncertainty (DIFF). 'EIGEN0' is the dominant variation of the eigen-variations.
FF_OTHER_SUBTRACTION	Uncertainty of other subtracted non-multi-jet processes
TF_STAT_1P_LEAD TF_STAT_1P_SUBL TF_STAT_3P_LEAD TF_STAT_3P_SUBL	Statistical uncertainty in the 1- or 3-prong *TF*s (fake τ_{had} is p_{T} leading (LEAD) or sub-leading (SUBL) τ_{had})

Finally, experimental uncertainties are considered, as encapsulated in Table 4.9 and 4.10. They are related to the responses of the ATLAS detector to the physics objects and are derived from dedicated measurements. These uncertainties are applied to all the simulated events. The origins of the uncertainties are shown in the descriptions. The uncertainty in the LHC Run2 integrated luminosity is 1.7%, obtained by the LUCID-2 detector [150]. The uncertainties in the trigger, reconstruction, identification, isolation efficiencies and the energy of electrons [151], muons [68] and τ_{had} [101] are estimated using tag-and-probe approach in analyses of $Z \to \ell\ell$, $J/\psi \to \ell\ell$ and $t\bar{t} \to b\mu\nu_\mu b\tau\nu_\tau$ events, where $\ell = e, \mu, \tau$. For muons and τ_{had} the uncertainties in track-to-vertex-association and electron-veto are also considered, respectively. The uncertainties related to the trigger, reconstruction, identification and isolation efficiencies of electrons and muons are only used in the $Z + hf$ CR with $b\bar{b}\ell^+\ell^-$ final state and the $\tau_{\text{lep}}\tau_{\text{had}}$ decay channel. Jet energy scale and resolution uncertainties [99] are estimated from the primary jet energy calibration based on simulation followed by an in situ correction based on data. The jet vertex tagging efficiency uncertainty [152] is measured using tag-and-probe method in $Z \to \mu\mu$+jets events. The uncertainties in b-tagging efficiency [79, 153, 154] are derived as functions of b-jet and c-jet p_{T} using $t\bar{t}$ events, and as functions of p_{T} and η of light-flavour jets using di-jet events.

4.4.2 Theoretical uncertainties

4.4.2.1 Theoretical uncertainty of signals

For the SM non-resonant HH signals, the uncertainties in the cross sections of productions and the branching ration of $H \to bb$ and $H \to \tau\tau$ decays are discussed in Section 1.2.2. For the resonant

Table 4.9: Summary of experimental systematic uncertainties (Part I). The uncertainties related to the trigger, reconstruction, identification and isolation efficiencies of electrons and muons are only used in the $Z + hf$ CR with $b\bar{b}\ell^+\ell^-$ final state and the $\tau_{\text{lep}}\tau_{\text{had}}$ decay channel.

Name of uncertainty	Description
LUMI_Run2	Uncertainty in Run2 integrated luminosity
PRW_DATASF	Uncertainty in reweighting pile-up distribution to data
EL_EFF_TRIG_TOTAL_1NPCOR_PLUS_UNCOR	Uncertainty in electron trigger efficiency
EL_EFF_RECO_TOTAL_1NPCOR_PLUS_UNCOR	Uncertainty in electron reconstruction efficiency
EL_EFF_ID_TOTAL_1NPCOR_PLUS_UNCOR	Uncertainty in electron identification efficiency
EL_EFF_ISO_TOTAL_1NPCOR_PLUS_UNCOR	Uncertainty in electron isolation efficiency
EG_RESOLUTION_ALL	Uncertainty in electron energy resolution
EG_SCALE_ALL	Uncertainty in electron energy scale
EG_SCALE_AF2	Uncertainty in electron energy scale for FastSim samples
MUON_EFF_TrigStatUncertainty	Uncertainties in muon trigger efficiency
MUON_EFF_TrigSystUncertainty	
MUON_EFF_RECO_STAT	Uncertainties in muon reconstruction efficiency
MUON_EFF_RECO_SYS	
MUON_EFF_RECO_STAT_LOWPT	
MUON_EFF_RECO_SYS_LOWPT	
MUON_EFF_ISO_STAT	Uncertainties in muon isolation efficiency
MUON_EFF_ISO_SYS	
MUON_EFF_TTVA_STAT	Uncertainties in muon track-to-vertex association (TTVA) efficiency
MUON_EFF_TTVA_SYS	
MUON_ID	Uncertainty in muon Inner Detector (ID) track resolution
MUON_MS	Uncertainty in muon Muon Spectrometer (MS) track resolution
MUON_SCALE	Uncertainty in muon energy scale
MUON_SAGITTA_RHO	Uncertainties in muon charge dependent energy scale
MUON_SAGITTA_RESBIAS	
TAUS_TRIGGER_STATDATA[16|17|18]	Uncertainties in τ_{had} trigger efficiency
TAUS_TRIGGER_STATMC[16|17|18]	
TAUS_TRIGGER_SYST[16|17|18]	
TAUS_TRIGGER_SYSTMU[16|17|18]	
TAUS_RECO_HIGHPT	Uncertainties in τ_{had} reconstruction efficiency
TAUS_RECO_TOTAL	
TAUS_RNNID_1PRONGSTATSYSTPT2025	Uncertainties in τ_{had} identification efficiency
TAUS_RNNID_1PRONGSTATSYSTPT2530	
TAUS_RNNID_1PRONGSTATSYSTPT3040	
TAUS_RNNID_1PRONGSTATSYSTPTGE40	
TAUS_RNNID_3PRONGSTATSYSTPT2025	
TAUS_RNNID_3PRONGSTATSYSTPT2530	
TAUS_RNNID_3PRONGSTATSYSTPT3040	
TAUS_RNNID_3PRONGSTATSYSTPTGE40	
TAUS_RNNID_HIGHPT	
TAUS_RNNID_SYST	
TAUS_ELEBDT_STAT	Uncertainties in τ_{had} electron-veto efficiency
TAUS_ELEBDT_SYST	
TAUS_ELEOLR_TOTAL	
TAUS_HADTAU_SME_TES_INSITUEXP	Uncertainties in τ_{had} energy scale
TAUS_HADTAU_SME_TES_INSITUFIT	
TAUS_HADTAU_SME_TES_MODEL_CLOSURE	
TAUS_HADTAU_SME_TES_PHYSICSLIST	
MET_SoftTrk_ResoPara	Uncertainty in the longitudinal energy resolution of the soft term of $E_{\text{T}}^{\text{miss}}$
MET_SoftTrk_ResoPerp	Uncertainty in the transverse energy resolution of the soft term of $E_{\text{T}}^{\text{miss}}$
MET_SoftTrk_Scale	Uncertainty in the energy scale of the soft term of $E_{\text{T}}^{\text{miss}}$

Table 4.10: Summary of experimental systematic uncertainties (Part II). The uncertainties related to the trigger, reconstruction, identification and isolation efficiencies of electrons and muons are only used in the $Z + hf$ CR with $b\bar{b}\ell^+\ell^-$ final state and the $\tau_{\text{lep}}\tau_{\text{had}}$ decay channel.

Name of uncertainty	Description
JET_EtaIntercalibration_Modelling	Uncertainties in jet energy scale
JET_EtaIntercalibration_TotalStat	
JET_EtaIntercalibration_NonClosure_highE	
JET_EtaIntercalibration_NonClosure_negEta	
JET_EtaIntercalibration_NonClosure_posEta	
JET_Pileup_OffsetMu	
JET_Pileup_OffsetNPV	
JET_Pileup_PtTerm	
JET_Pileup_RhoTopology	
JET_Flavor_Composition	
JET_Flavor_Response	
JET_PunchThrough_MC16	
JET_PunchThrough_AFII	
JET_EffectiveNP_Detector1	
JET_EffectiveNP_Detector2	
JET_EffectiveNP_Mixed1	
JET_EffectiveNP_Mixed2	
JET_EffectiveNP_Mixed3	
JET_EffectiveNP_Modelling1	
JET_EffectiveNP_Modelling2	
JET_EffectiveNP_Modelling3	
JET_EffectiveNP_Modelling4	
JET_EffectiveNP_Statistical1	
JET_EffectiveNP_Statistical2	
JET_EffectiveNP_Statistical3	
JET_EffectiveNP_Statistical4	
JET_EffectiveNP_Statistical5	
JET_EffectiveNP_Statistical6	
JET_SingleParticle_HighPt	
JET_RelativeNonClosure_AFII	
JET_BJES_Response	
JET_EtaIntercalibration_NonClosure_2018data	
JET_JER_DataVsMC_MC16	Uncertainties in jet energy resolution
JET_JER_DataVsMC_AFII	
JET_JER_EffectiveNP_1	
JET_JER_EffectiveNP_2	
JET_JER_EffectiveNP_3	
JET_JER_EffectiveNP_4	
JET_JER_EffectiveNP_5	
JET_JER_EffectiveNP_6	
JET_JER_EffectiveNP_7	
JET_JER_EffectiveNP_8	
JET_JER_EffectiveNP_9	
JET_JER_EffectiveNP_10	
JET_JER_EffectiveNP_11	
JET_JER_EffectiveNP_12restTerm	
JET_JVT_EFF	Uncertainty in jet vertex tagging efficiency
JET_FJVT_EFF	
FT_EFF_Eigen_B_[0-2]	Uncertainties in b-tagging efficiency for b-jets
FT_EFF_Eigen_C_[0-3]	Uncertainties in b-tagging efficiency for c-jets
FT_EFF_Eigen_Light_[0-4]	Uncertainties in b-tagging efficiency for light flavour (u, d, s) jets
FT_EFF_extrapolation	Uncertainties in b-tagging efficiency for high-p_{T} jets
FT_EFF_extrapolation_from_charm	Uncertainties in b-tagging efficiency for jets initiated from τ_{had}

HH signals, these uncertainties are not considered because they are BSM processes. However, an additional uncertainty to account for the difference between FastSim and FullSim samples is applied as mentioned in Section 4.1. It is a 6.5% overall uncertainty.

Commonly for non-resonant and resonant HH signals, the uncertainties in the parton shower and hadronisation (PS uncertainty) are estimated by comparing the nominal samples generated by PYTHIA 8 (HERWIG 7 for resonant signals) to the alternative samples generated by HERWIG 7 (PYTHIA 8 for resonant signals). They are found to be the dominant source of uncertainty and only have normalisation effect in all cases. For resonant signal the magnitude of the uncertainties is estimated using several exemplary mass points and interpolated to the other mass points in between. Other sources of uncertainties are found to be negligible, the following describes the procedure to derive them. The uncertainties in renormalisation and factorisation scales (SCALE uncertainty) are estimated by varying the values μ_R and μ_F by a factor of 2 or 0.5 in 7 points

$$(\mu_R, \mu_F) \rightarrow (0.5, 0.5), (0.5, 1), (1, 0.5), (1, 1), (1, 2), (2, 1), (2, 2). \tag{4.12}$$

The maximum variation among the 7-point variations represents the final uncertainty. PDF4LHC is used for these samples. The uncertainties of PDF and α_s (PDFalphas) for the signals are estimated in the approach recommended by Ref. [133], where total PDF+α_s uncertainty is estimated.

Table 4.11 summarises the $\mathcal{A} \times \epsilon$ uncertainties for the simulated signal samples in the SR. The relative uncertainties are presented. Non-resonant SM ggF and VBF signals and resonant signals with mass hypotheses of 500 and 1000 GeV are shown as examples.

Table 4.11: Summary of the $\mathcal{A} \times \epsilon$ uncertainties of the simulated HH signal samples in the SR. The numbers in the table are the relative uncertainties of the source on that column and the process in that row. The abbreviations for source of uncertainties are explained in the text.

Process	PS	SCALE	PDFalphas
SM ggF HH	4.3%	1.4%	0.1%
SM VBF HH	3.0%	<0.1%	1.0%
$X \rightarrow HH(m_X = 500\text{GeV})$	6.0%	<0.1%	<0.1%
$X \rightarrow HH(m_X = 1000\text{GeV})$	5.3%	<0.1%	<0.1%

4.4.2.2 Theoretical uncertainty of $Z + hf$ background

The simulation of Z+jets is studied in Ref. [143, 155]. Large discrepancies are found between the nominal samples generated by SHERPA 2.2 and the alternative samples generated by MADGRAPH5_aMC@NLO + PYTHIA 8 for processes involving heavy flavour jets in Ref. [143]. Thus, an uncertainty of the comparison of predictions made by different generators is introduced as an uncertainty (GENERATOR uncertainty). As discussed in Section 4.4.1, the relative normalisation uncertainty on $\mathcal{A} \times \epsilon$ and the shape uncertainty are derived separately in the SR. The normalisation uncertainty, together with those from other sources, is summarised in Table 4.12. It is found

that the shape uncertainty can be linearly parameterized using m_{bb} as shown in Figure 4.21(a) and propagated to the MVA output scores.

Similarly as for the signals, the SCALE and PDFalphas uncertainties are considered. The variations of μ_R and μ_F affect the shape of the distributions of MVA score outputs. Following the same '7-point' approach, the shape variation is estimated by treating the largest shape variation in MVA score distribution (using the binning for likelihood fit) as shape uncertainty, as illustrated in Figure 4.21(b)– 4.21(d) for several MVA scores. For $Z + hf$ baseline samples, the NNPDF3.0NNLO is used as the PDF model, the uncertainties are derived following Ref. [134]. In addition, the samples generated using alternative PDFs (MMHT20142014NNLO68CL and CT14NNLO) are compared with the baseline sample. The differences are treated as PDFChoice uncertainty to account for the choice of PDF. This uncertainty can be ignored if it is significantly smaller compared with the PDFalphas uncertainty.

Two more sources of uncertainties are required for V+jets as recommended in Ref. [155]. The first one is CKKW uncertainty which accounts for the configuration of CKKW [156] merging scale used in the calculation for double counted jets from ME calculated at NNLO precision and PS. The second one is QSF uncertainty which relates to the resummation scale used in the PS model to begin the process of soft emission.

The relative normalisation uncertainties in the SR with respect to the $Z+hf$ CR are summarised in Table 4.12. The ones that are correlated with shape uncertainties are marked with 'Has shape'. The shape variations in the distributions of various MVA scores are illustrated in Figure 4.22.

Table 4.12: Summary of the relative normalisation uncertainties on $\mathcal{A} \times \epsilon$ of the simulated $Z + hf$ samples in the SR with respect to the $Z + hf$ CR. The abbreviations for source of uncertainties are explained in the text.

Source of uncertainty	Relative Uncertainty	Has shape
GENERATOR	7.0%	√
SCALE	$^{+12\%}_{-9.6\%}$	√
PDFChoice	0.98%	
PDFalphas	0.77%	
CKKW	5.3%	
QSF	6.0%	

4.4.2.3 Theoretical uncertainty of $t\bar{t}$ background

The simulation of $t\bar{t}$ process is presented in [157]. Theoretical uncertainties are estimated for the true-τ_{had} $t\bar{t}$ simulated samples. For fake-τ_{had} $t\bar{t}$ samples, the theoretical uncertainties are not applied as they are corrected by data-driven scale factors.

To estimate the uncertainties arise from different approaches to the ME calculated at NLO precision and PS matching (ME uncertainty), the POWHEG BOX V2 +PYTHIA 8 nominal samples are compared with the MADGRAPH5_aMC@NLO +PYTHIA 8 alternative samples [158]. The uncertainty of

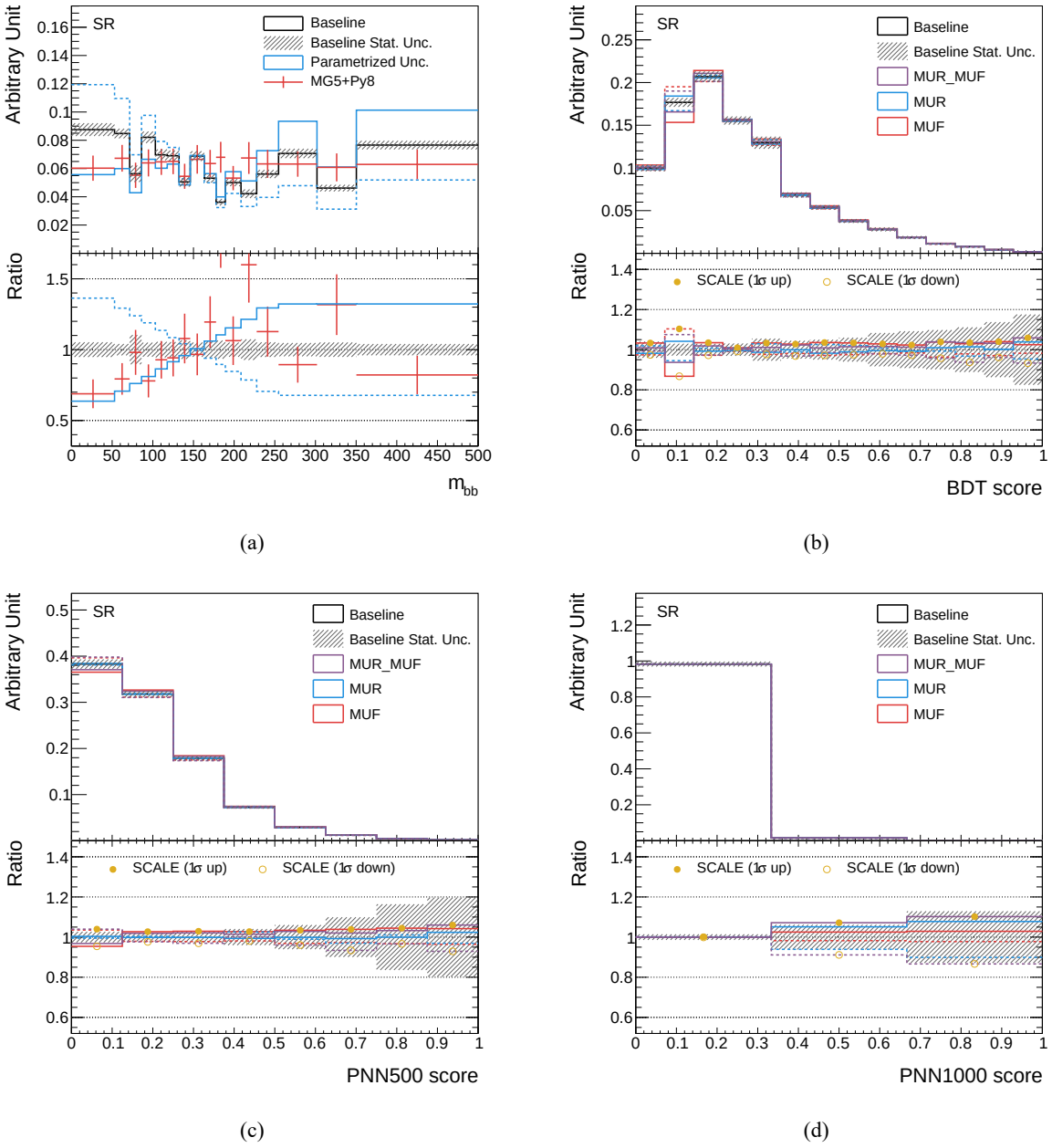


Figure 4.21: (a) Shape only comparison of the nominal and alternative $Z+hf$ samples, the discrepancy is parameterized as a piecewise linear function shown in cyan line. (b,c,d) Shape variations of renormalisation and factorisation scale uncertainties estimated in comparison of '7-point' variations. For each MVA score distribution, the largest shape variation is considered as the shape uncertainty. Three MVA scores are shown; non-resonant HH BDT (left), resonant HH PNN with $m_X = 500$ GeV (middle) and $m_X = 1000$ GeV (right).

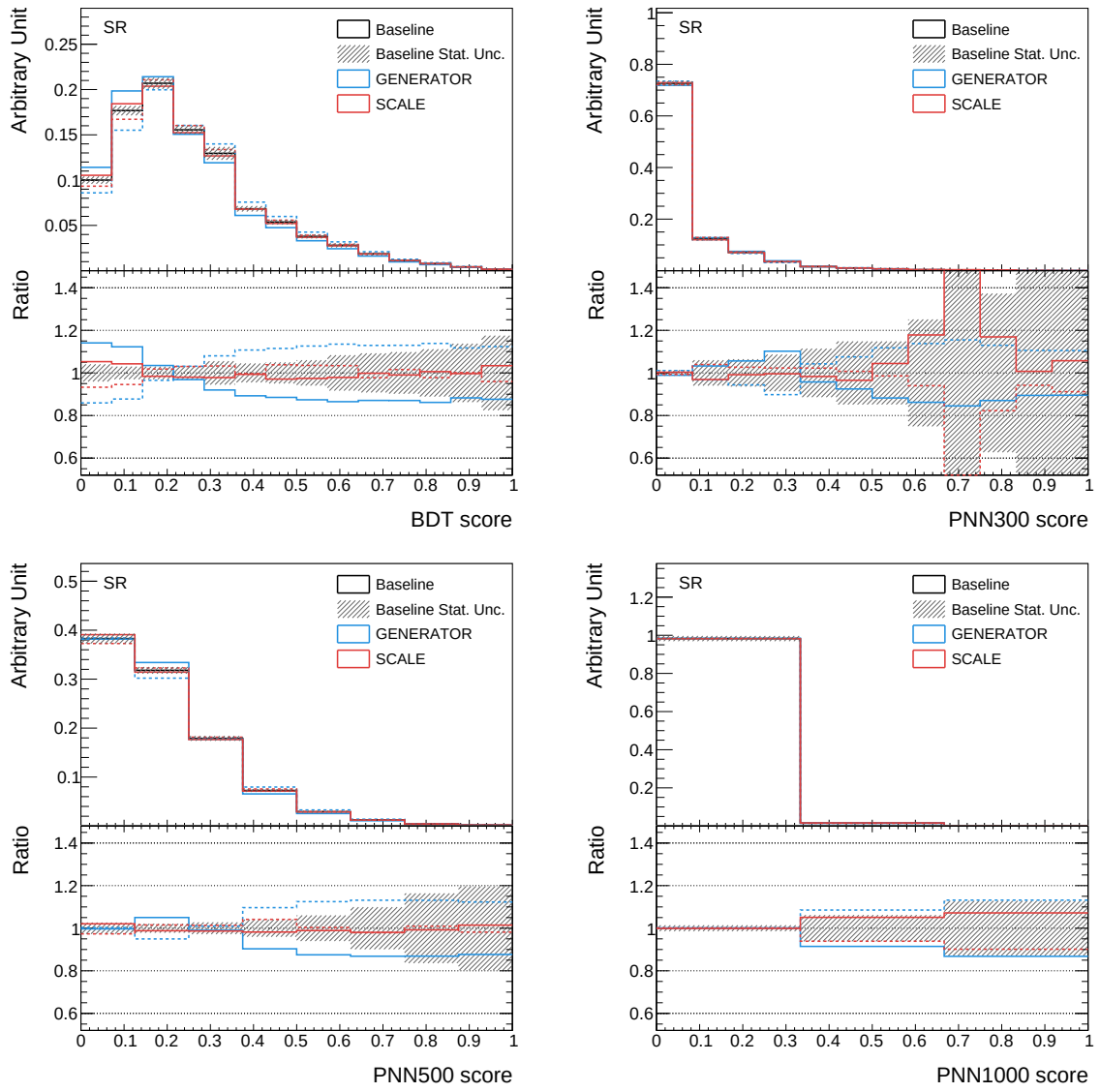


Figure 4.22: Shape variations of the uncertainties on non-resonant HH BDT, resonant HH PNN with $m_X = 300$, 500 and 1000 GeV for the simulated $Z + hf$ samples in SR. The distributions are normalised to unit integral to visualise the shape variations. The hatched areas are the statistical uncertainty of the baseline the simulated $Z + hf$ samples. Solid and dashed lines stand for 1σ up and down variations.

the h_{damp} parameter in POWHEG BOX V2 is estimated by varying its nominal value ($1.5m_{\text{top-quark}}$) by a factor of 2 and 0.5. The h_{damp} parameter is used to regulate the first quark/gluon emission of the ME generator. The estimation of the PS uncertainties is the same as that for the signal processes. The uncertainty in the tuning configuration of the initial-state radiation of the PS generator (ISR uncertainty), as well as on the tuning configuration of the final-state radiation of the PS generator (FSR uncertainty) are estimated by varying the corresponding parameters in PYTHIA 8. The SCALE uncertainty is estimated jointly with the ISR uncertainty. The PDFalphas uncertainty is estimated by the same procedure as that for the signal processes.

Table 4.13 outlines the relative in the SR with respect to the $Z + hf$ CR. The ones that are correlated with shape uncertainties are marked in 'Has shape', whose shape variations on the distributions of various MVA scores are illustrated in Figure 4.23.

Table 4.13: Summary of the relative normalisation uncertainties on $\mathscr{A} \times \epsilon$ of $t\bar{t}$ simulated samples in the SR with respect to the $Z+hf$ CR. The abbreviations for source of uncertainties are explained in the text.

Source of uncertainty	Relative Uncertainty	Has shape
ME	3.7%	√
PS	2.2%	√
ISR	0.30%	√
FSR	1.9%	√
PDFalphas	0.011%	

4.4.2.4 Theoretical uncertainty of other backgrounds

The other background simulated samples in consideration of theoretical uncertainties are single top-quark, single Higgs boson, $Z+lf$, W+jets, VV. Uncertainties on the cross sections for normalisation are considered for all these processes, as listed in Table 4.14. Single top-quark and single Higgs boson, the uncertainties on $\mathscr{A} \times \epsilon$ are estimated in the SR. Branching ratio uncertainties on $H \to bb$ and $H \to \tau\tau$ are considered for the single Higgs boson processes.

Same generators are used to generate the single top-quark and top-quark pair processes. Hence, the $\mathscr{A} \times \epsilon$ uncertainties are derived using the same approach. They are only derived for the Wt production channel as the contributions of s-channel and t-channel is negligible in the SR. According to Ref. [159], an additional uncertainty needs to be considered for the methods that are used to calculate the interference between the LO diagram of $t\bar{t}$ and the NLO diagram of the Wt-channel of single top-quark in the generator (Top_Interference uncertainty). The recommended way is to estimate the uncertainty by the difference between the Diagram Removal (DR) and Diagram Subtraction (DS) methods [160]. This uncertainty is the dominant uncertainty for single top-quark modelling. It is found that only Top_Interference uncertainty has shape effects on the MVA scores, and it can be parameterized by the transverse momentum of bb system p_{T}^{bb} and propa-

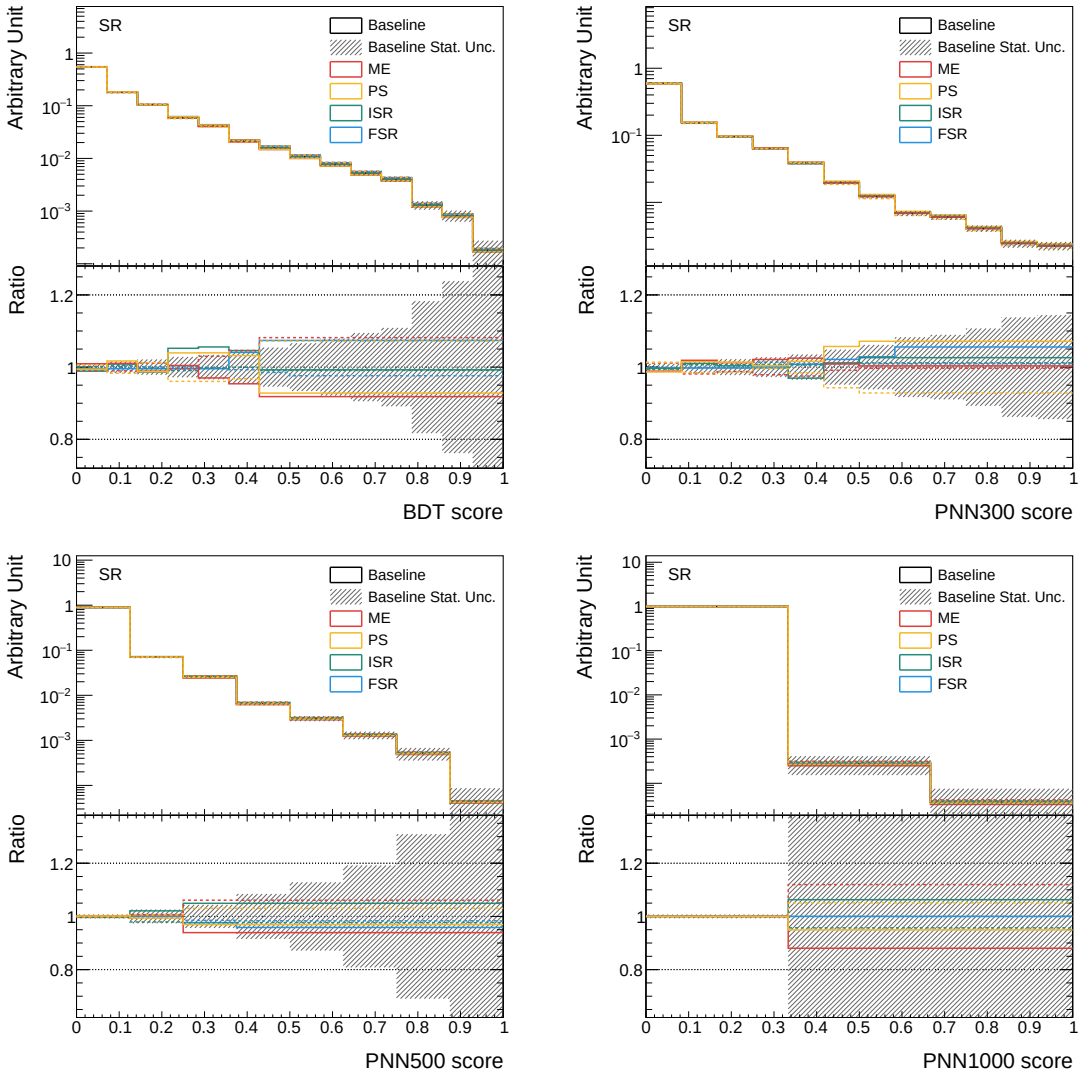


Figure 4.23: Shape variations of the uncertainties on non-resonant HH BDT, resonant HH PNN with $m_X = 300$, 500 and 1000 GeV for $t\bar{t}$ simulated samples in SR. The distributions are normalised to unit integral to visualise the shape variations. The hatched areas are the statistical uncertainty of the baseline $t\bar{t}$ simulated samples. Solid and dashed lines stand for 1σ up and down variations.

Table 4.14: Uncertainties on the cross sections for normalisation are considered for various processes.

Process	Relative Uncertainty
Single top-quark (Wt-channel)	5.4%
Single top-quark (s-channel)	3.7%
Single top-quark (t-channel)	4.2%
Single Higgs boson (ggF)	$\pm$3.2%(PDF and α_S) $\pm$ 3.9(Scale)
Single Higgs boson (VBF)	$\pm$2.1%(PDF and α_S)$^{+0.4\%}_{-0.3\%}$(Scale)
Single Higgs boson ($qq \to ZH$)	$\pm$1.9%(PDF and α_S)$^{+0.5\%}_{-0.6\%}$(Scale)
Single Higgs boson ($gg \to ZH$)	$\pm$2.4%(PDF and α_S)$^{+25\%}_{-19\%}$(Scale)
Single Higgs boson (WH)	$\pm$1.9%(PDF and α_S)$^{+0.5\%}_{-0.7\%}$(Scale)
Single Higgs boson ($t\bar{t}H$)	$\pm$3.6%(PDF and α_S)$^{+5.8\%}_{-9.2\%}$(Scale)
$Z + lf$	5.0%
W+jets	5.0%
VV	6.0%

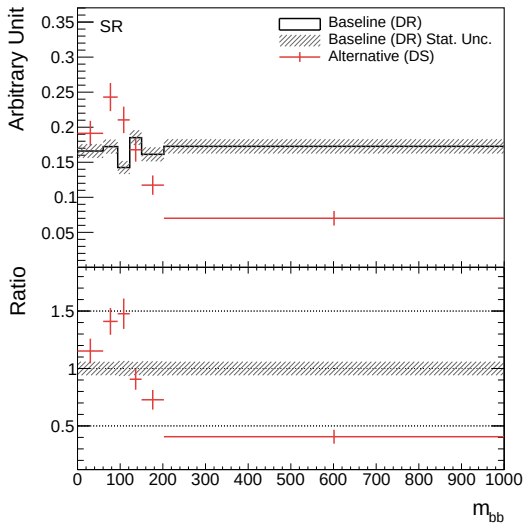
gated to the MVA scores. The parametrization and shape variation of the `Top_Interference` uncertainty on MVA scores are shown in Figure 4.24. Normalisation uncertainties are summarised in Table 4.15.

Table 4.15: Summary of the normalisation uncertainties on $\mathscr{A} \times \epsilon$ of the Wt production channel of single top-quark simulated samples in the SR. The abbreviations for source of uncertainties are explained in the text.

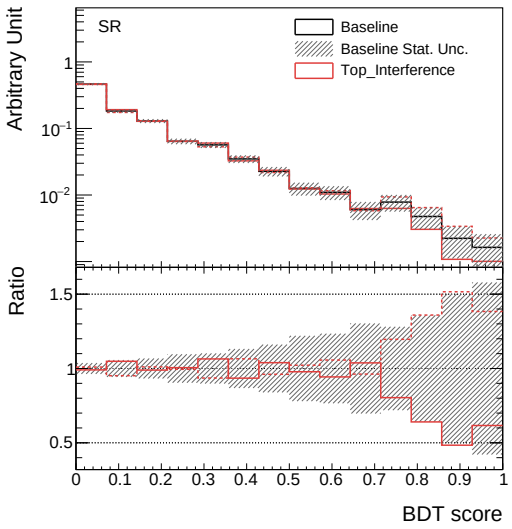
Source of uncertainty	Relative Uncertainty	Has shape
`Top_Interference`	27%	$\surd$
`ME`	4.9%	
`PS`	16%	
`ISR`	$^{+8.7\%}_{-9.4\%}$	
`FSR`	$^{+6.6\%}_{-4.9\%}$	
`PDFalphas`	3.2%	

For the single Higgs boson processes that have major contribution to the SR, the $\mathscr{A} \times \epsilon$ uncertainties are evaluated in the similar way as above. These processes include ZH and $t\bar{t}H$. Besides, to account for the difficulty in the modelling of single Higgs boson processes produced with heavy-flavour jets initiated from radiations [161, 162], a 100% uncertainty is applied to ggF, VBF and WH single Higgs productions where the Higgs boson decays to τ-leptons (`HF` uncertainty). All the $\mathscr{A} \times \epsilon$ uncertainties are found to only have normalisation effect, the magnitudes are summarised in Table 4.16.

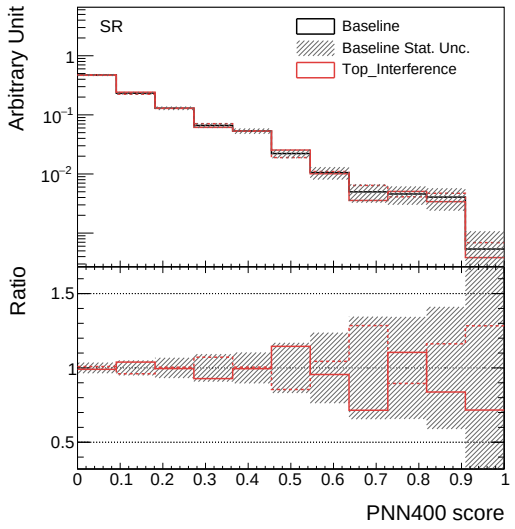
For the rest of the processes, their contributions are so minor that are not feasible to estimated in the $b\bar{b}\tau^+\tau^-$ SR, instead, the uncertainties are learned from the estimation in $VH(H \to bb)$ analysis [163]. The values are listed in Table 4.17, only normalisation effect is considered.



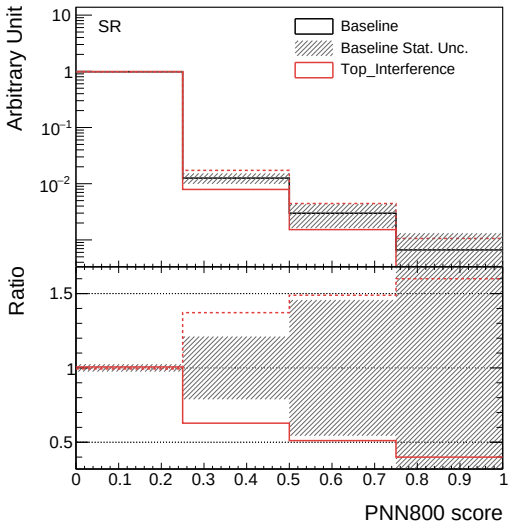
(a) Parametrization



(b) Variation on Non-resonant HH BDT



(c) Variation on PNN at $m_X = 400$ GeV



(d) Variation on PNN at $m_X = 800$ GeV

Figure 4.24: Parameterization of the `Top_Interference` uncertainty in the Wt production channel of single top-quark simulated samples in the SR using p_{T}^{bb} variable (a). The parameterized uncertainty is propagated to Non-resonant HH BDT and PNN at various m_Xs in (b), (c) and (d).

Table 4.16: Summary of the normalisation uncertainties on $\mathscr{A} \times \epsilon$ of the single Higgs boson simulated samples in the SR. The abbreviations for source of uncertainties are explained in the text.

Process	Source	Relative Uncertainty
ggF, VBF, WH with $H \to \tau\tau$	HF	100%
$ZH, H \to \tau\tau$	PS	5.6%
	SCALE	2.6%
	PDFalphas	1.0%
$ZH, H \to bb$	PS	12%
	SCALE	1.8%
	PDFalphas	1.0%
$t\bar{t}H$	ME	3.3%
	PS	1.1%
	FSR	$^{+3.7\%}_{-6.0\%}$
	SCALE	1.0%

Table 4.17: Summary of the normalisation uncertainties on $\mathscr{A} \times \epsilon$ of the simulated samples of various minor background processes in the SR. The values are from the estimation in $VH(H \to bb)$ analysis [163]. The abbreviations for source of uncertainties are explained in the text.

Process	Relative Uncertainty
Single top-quark (s-channel)	20%
Single top-quark (t-channel)	20%
W+jets	50%
$Z + lf$	23%
WW	25%
ZZ	26%
WZ	20%

4.5 Statistical analysis

Distributions of the output scores of MVA algorithms are used in the procedure in determining the signal strength and performing other statistical interpretations, as the high MVA score regions efficiently locate the areas that are sensitive to the presence of signal processes. The MVA score distributions are binned following the descriptions in Section 4.2.4.4 such that the number of bins and statistics in each bin are controlled to ensure numerical stability of the minimization programs used in maximum likelihood parameter estimation and the validity of the asymptotic approximation (see Section 4.5.2), and simultaneously, to achieve close to optimal sensitivity of searching signals. The MVA score distributions in the chosen binnings are demonstrated in Figure 4.25. The $m_{\ell\ell}$ distribution of the $Z + hf$ CR is used to control the normalisation of $Z + hf$ and $t\bar{t}$ background, as discussed in Section 4.3.1. It is combined with the SR in the analysis.

Each bin of the distributions are treated as counting experiments that are described by Poisson random processes. The (nuisance) parameters are constrained by random distributions that relate to the auxiliary measurements on the corresponding systematic uncertainties. The probability density functions ($p.d.f$s) are built based on the predictions of the signal and background components simulated as discussed in Section 4.1 and Section 4.3, and the uncertainties in Section 4.4. Consequently, the likelihood model is constructed by the product of the $p.d.f$s.

4.5.1 Likelihood function

According to the introduction above, the likelihood function constructed from the SR and $Z + hf$ CR can be schematically written as

$$\mathscr{L}(\mu, \boldsymbol{\theta}) = \prod_i p_i(n_i; \mu, \boldsymbol{\theta}) \prod_a p_a^{\text{auxiliary}}(m_a; \theta_a), \tag{4.13}$$

where i is the index of the counting experiment in the i-th bin and a is the index of the auxiliary measurement on the systematic uncertainty related to the a-th nuisance parameter. Each ingredient in Equation (4.13) is elaborated in the following.

The likelihood $\mathscr{L}$ is in terms of parameters μ and $\boldsymbol{\theta}$, where μ is the parameter of interest (POI) that measures the normalisation of the magnitude of signal known as the signal strength, and $\boldsymbol{\theta}$ is the nuisance parameters (NPs) that parametrise the impacts of the systematic uncertainties. The set of NPs consist of $\boldsymbol{\mu}_\text{b}$ which relates to the normalisation of the background processes, $\boldsymbol{\gamma}$ which relates to the statistical uncertainties of the simulated samples and $\boldsymbol{\alpha}$ which relates to the other systematic uncertainties, i.e.

$$\boldsymbol{\theta} = \{\boldsymbol{\mu}_\text{b}, \boldsymbol{\gamma}, \boldsymbol{\alpha}\}. \tag{4.14}$$

More details about their roles in the likelihood function are introduced in the next paragraphs. The nuisance parameters can be profiled out by maximising $\mathscr{L}(\mu, \boldsymbol{\theta})$ with a fixed value of μ. The

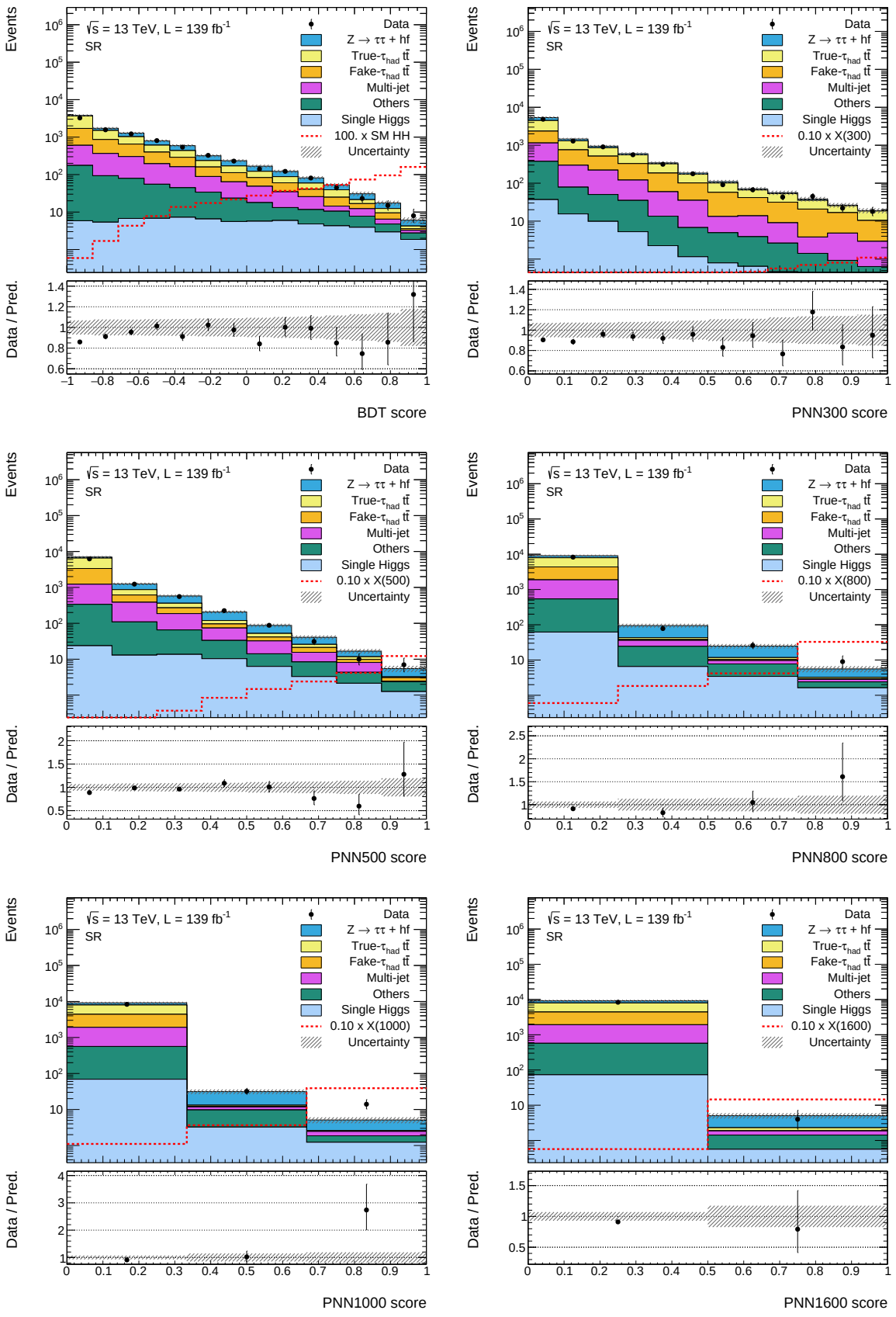


Figure 4.25: Distributions of the BDT score for the non-resonant HH search and PNN evaluated at $m_X = 300,500,800,1000,1600$, by which the likelihood functions are constructed. The cross section of non-resonant HH signal is scaled to the SM value while those of the resonant $X \to HH$ signals are scaled to 1 pb, and they are further scaled by the numbers shown in the legend. In the lower panel, the ratio of data and the total background prediction is displayed in black dots. The hatched area indicates the total systematic uncertainty of the backgrounds.

maximum likelihood estimators of $\boldsymbol{\theta}$ given a fixed μ can be written as

$$\hat{\boldsymbol{\theta}}_{\mu} = \arg\max_{\boldsymbol{\theta}} \mathscr{L}_{\mu}(\boldsymbol{\theta}), \tag{4.15}$$

where μ in the subscripts indicates that the value of μ is fixed in the evaluation of the estimators. This results in a likelihood in the form of $\mathscr{L}(\mu, \hat{\boldsymbol{\theta}}_{\mu})$, which is called a profile likelihood. It is a function only about the parameter μ that are interesting to our analysis. A mathematical review on the principles of profile likelihood functions can be found in Ref. [164].

In the construction of the likelihood function, the $p.d.f$s are separated into two categories to illustrate that they are defined by different approaches in an experimental point of view. In the first part, p_i describes the probability distribution of the counting experiment in the i-th bin. It follows the Poisson p.d.f[17],

$$p_i(n_i; \mu, \{\boldsymbol{\mu}_b, \boldsymbol{\gamma}, \boldsymbol{\alpha}\}) = Poisson\left(n_i; \gamma_i\left[\mu S_i(\boldsymbol{\alpha}) + \sum_b \mu_b B_{bi}(\boldsymbol{\alpha})\right]\right), \tag{4.16}$$

where n_i is the random number of events; S_i stands for the expected number of signal events; B_{bi} stands for the expected number of background events of b process [18]. μ_b is the normalisation factor of background b, which only exists if b is $Z + hf$ or $t\bar{t}$ [19]. γ_i is motivated by the Beeston-Barlow technique [165] that provides an approach to consider the statistical uncertainties of the simulated samples in the likelihood function. The original Beeston-Barlow method assigns a γ for each process and bin. A simplified version is used in HistFactory where γ is assigned per bin to account for the total statistical uncertainty of all processes. Note that γ_i is not applied to the signals because the statistical uncertainties on signals are known to be negligible.

In the second part, the probability distribution of the auxiliary experiments related to the a-th nuisance parameter θ_a is described by $p_a^{\text{auxiliary}}$. The auxiliary experiments are the dedicated measurements that estimate the effect of the systematic uncertainties, which can be parameterized by the corresponding nuisance parameters such that the effect can be interpolated from the estimated case (considered as $\pm 1\sigma$ variation) to the nominal case (no variation). Default interpolation algorithms described in [166] are employed.

Strictly, the full probability distributions of the auxiliary experiments, or in other words, the full likelihood functions of the nuisance parameters need to be included. However, this is usually unworkable because there are many auxiliary experiments, for instance, to measure detector response in ATLAS. As a good approximation, one can assume that if the auxiliary experiment is repeated many times, the maximum likelihood estimators (the measured value) of the nuisance parameters will follow a certain distribution $p_a^{\text{auxiliary}}(m_a; \theta_a)$. The distribution measures the probability distribution of the randomly fluctuated measurement of θ_a, denoting as m_a. The best-fit value $\hat{\theta}_a$ of

[17] $Poisson(n; \nu) = \frac{\nu^n}{n!} e^{-\nu}$

[18] e.g. $b = Z + hf$

[19] The same parameter is applied on true-τ_{had} and fake-τ_{had} $t\bar{t}$ processes.

auxiliary experiment is considered to be the true value of θ_a.

Two kinds of auxiliary experiments are involved in the likelihood function in the analyses of this thesis. The same concept is widely used in ATLAS physical analyses. A detailed explanation is provided below. The first one measures the statistical uncertainty of the simulated samples. If the $\pm1\sigma$ relative statistical uncertainty in the i-th bin is δ_i, defining $\tau_i = 1/\delta_i$ and considering the generations of samples as auxiliary experiments, the measurement of τ_i^2 – random variable m_i should follow a Poisson distribution

$$Poisson(m_i; \gamma_i\tau_i), \tag{4.17}$$

where γ_i is the nuisance parameter introduced by Beeston-Barlow method that parametrise the uncertainty, whose true value is 1; τ_i is a constant that equals to the inverse of $\pm1\sigma$ relative statistical uncertainty, i.e., $\tau_i = 1/\delta_i$. Equation (4.17) is called a Poisson constrain term in the sense that the auxiliary measurement constrains the estimation of γ.

The second one relates to the other source of systematic uncertainties, whose effect is parameterized by $\boldsymbol{\alpha}$. For instance, in an auxiliary measurement of observable $\mathcal{O}$, $\delta(\alpha)$ are the measured relative uncertainties parameterized by α (by some interpolation algorithm), where $\alpha = \pm1$ corresponds to the $\pm1\sigma$ variation δ and $|\delta| = |\delta(\alpha = 1)|$ [20]. Therefore, the true value $\mathcal{O}$ can be parameterized by α

$$\mathcal{O}(\alpha) = (1 + \delta(\alpha))\mathcal{O}_c, \tag{4.18}$$

where $\mathcal{O}_c$ is a constant. So the 1σ variance is $\delta\mathcal{O}_c$ and the measured value of $\mathcal{O}(\alpha)$ is $\mathcal{O}(\alpha_m)$, where α_m denotes the measured value of α. It is a good approximation that by repeating the auxiliary measurement many times, $\mathcal{O}(\alpha_m)$ follows Gaussian distribution

$$Gaussian(\mathcal{O}(\alpha_m); \mathcal{O}(\alpha), \delta\mathcal{O}_c) = \frac{1}{\sqrt{2\pi(\delta\mathcal{O}_c)^2}} \exp\left[-\frac{[\mathcal{O}(\alpha_m) - \mathcal{O}(\alpha)]^2}{2(\delta\mathcal{O}_c)^2}\right]. \tag{4.19}$$

Furthermore, $\delta \cdot \alpha$ is a good approximation of $\delta(\alpha)$ for the piecewise linear and exponential interpolations that are used. It is valid to scale the observable by a factor of $1/\delta\mathcal{O}_c$ and substitute $\mathcal{O}(\alpha)$ with Equation (4.18) and the approximation $\delta(\alpha) = \delta \cdot \alpha$, i.e.,

$$\begin{aligned} \mathcal{O}_c &\xrightarrow{\text{Scale}} 1/\delta \\ \mathcal{O}_m &\xrightarrow{\text{Scale}} 1/\delta + \alpha_m \\ \mathcal{O} &\xrightarrow{\text{Scale}} 1/\delta + \alpha. \end{aligned} \tag{4.20}$$

[20] The variation is assume to be symmetric in the $\pm1\sigma$ directions in this example. The interpolation can handle asymmetric cases and this example can be generalised to that case.

Consequently, after the scaling, Equation (4.19) can be written as

$$Gaussian(\mathcal{O}(\alpha_m); \mathcal{O}(\alpha), \delta\mathcal{O}_c) = \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(\alpha_m - \alpha)^2}{2}\right], \tag{4.21}$$

which can be denoted as $Gaussian(\alpha_m; \alpha, 1)$, known as the Gaussian constrain term.

In summary, by integrating all the ingredients into Equation (4.13), the likelihood function can be expressed as

$$\begin{aligned} \mathscr{L}(\mu, \{\boldsymbol{\mu_b}, \boldsymbol{\gamma}, \boldsymbol{\alpha}\}) = &\prod_i Poisson\left(n_i; \gamma_i\left[\mu S_i(\boldsymbol{\alpha}) + \sum_b \mu_b B_{bi}(\boldsymbol{\alpha})\right]\right) \\ &\times \prod_i Poisson(m_i; \gamma_i \tau_i) \\ &\times \prod_a Gaussian(\alpha_{m,a}; \alpha_a, 1), \end{aligned} \tag{4.22}$$

where the functional forms of $S_i(\boldsymbol{\alpha})$ and $B_{bi}(\boldsymbol{\alpha})$ are determined by the interpolation algorithms in use.

The software implementation of the likelihood function is powered by HistFactory [166]. The function minimisation in the likelihood fit (maximum likelihood estimation) is executed by MINUIT [167], which provides the best-fit values and confidence intervals for the parameters.

4.5.2 Statistical interpretation

The result of a search for new phenomena can be statistically interpreted from two aspects. One of them quantifies the significance level (Z) of observing the new phenomenon. Translating into the terminology of frequentist hypothesis test, in the context of the analyses in this thesis, the null hypothesis (H_0) where only background is expected is tested against the alternative hypothesis (H_A) where a certain amount of signal is expected, measured by signal strength μ, i.e.,

$$\begin{aligned} H_0 &: \mu = 0 \\ H_A &: \mu = \mu_A, \end{aligned} \tag{4.23}$$

where μ_A is the hypothetical (positive) signal strength by H_A. Rejecting the background-only hypothesis is one of the most recognizable standard to claim the discovery of a new phenomenon, which corresponds to a p-value of $2.87 \cdot 10^{-7}$ and a significance level Z of 5, known as the 5σ metric. Here p-value is converted to Z by the quantile of a normal distribution, $Z = \Phi^{-1}(1-p)$, in which Φ is the cumulative distribution function (c.d.f.) of the normal distribution. The p-value and significance derived in this way is dedicated to a single signal point. In the case that the searching space is a range, they are named as local p-value (p^{local}) and local significance (Z^{local}). To consider the probability of seeing an excess anywhere in the search range, global p-value ($p^{\text{glob.}}$) need to be

calculated, and the corresponding significance is called global significance ($Z^{\text{glob.}}$). Since a wide range of signal space is considered, $p^{\text{glob.}}$ is always higher than p^{local} (i.e. $Z^{\text{glob.}} < Z^{\text{local}}$). This is known as look elsewhere effect [168]. The global p-value is defined by the p.d.f of the maximum local significance ($f(Z^{\text{local}}_{\text{max}})$) of the full search range

$$p^{\text{glob.}} = \int_{Z^{\text{local}}_{\text{max,obs.}}}^{+\infty} f(Z^{\text{local}}_{\text{max}})\, dZ^{\text{local}}_{\text{max}}, \tag{4.24}$$

where $Z^{\text{local}}_{\text{max,obs.}}$ is the observed maximum local significance.

The other one aims to set an upper limit on the parameter of interest, which is μ in this context. In this case, the null hypothesis expects a certain amount of signal, which is tested against the alternative hypothesis that expects less signal. It can be expressed as

$$\begin{aligned} H_0 &: \mu = \mu_0 \\ H_A &: \mu < \mu_0. \end{aligned} \tag{4.25}$$

To find the upper limit of μ at a specified confidence-level, which is often set to 95% by ATLAS convention, we scan various null hypotheses H_0 with various values of μ_0 to find a upper bound μ_0^{up}, such that for all H_0 with μ_0 less than μ_0^{up}, the p-value calculated by the CL_s technique [169] satisfies $p_s < 1 - CL_s$. The CL_s technique provides a way to interpret the problem as a test on the signal-only null hypothesis, instead of a test of the background+signal hypothesis. The latter causes issue when the background contamination is high. Suppose the test statistic of the problem is t, and its distributions under signal+background and background-only hypotheses are t_{s+b} and t_b, respectively. Given an observed t, denoted as t_{obs}, the corresponding p-values are

$$\begin{aligned} p_{s+b} &= \int_{t_{\text{obs}}}^{\infty} t_{s+b} dt \equiv CL_{s+b} \\ p_b &= 1 - \int_{t_{\text{obs}}}^{\infty} t_b dt \equiv 1 - CL_b. \end{aligned} \tag{4.26}$$

To approach a test on signal-only, CL_s is defined by the conditional probability

$$\begin{aligned} p_s &= P(t_{s+b} \geq t_{\text{obs}} | t_b \geq t_{\text{obs}}) \\ &= p_{s+b}/(1 - p_b) \\ &= \frac{CL_{s+b}}{CL_b} \equiv CL_s. \end{aligned} \tag{4.27}$$

The profile likelihood ratio test statistic is used, which generalises the application of Neyman-

Pearson lemma to the case in which nuisance parameters are involved. It is

$$t_\mu = -2\ln\lambda_\mu = -2\ln\frac{\mathscr{L}(\mu,\hat{\hat{\boldsymbol{\theta}}})}{\mathscr{L}(\hat{\mu},\hat{\boldsymbol{\theta}})}, \tag{4.28}$$

where the numerator inside the logarithm is the profile likelihood in which $\hat{\hat{\boldsymbol{\theta}}}$ is equivalent to the $\hat{\boldsymbol{\theta}}_\mu$ that was introduced in Equation (4.15) [21]; the denominator is the maximum likelihood in which $\hat{\mu}$ and $\hat{\boldsymbol{\theta}}$ are the estimators of μ and $\boldsymbol{\theta}$, respectively. By definition, larger value of t_μ stands for stronger inconsistency between the prediction of the hypothesis and the observed data. Considering that the signal has a positive contribution, t_μ is defined in the following form, denoting as $\tilde{t}_\mu$ to distinguish from the common case,

$$\tilde{t}_\mu = -2\ln\tilde{\lambda}_\mu = \begin{cases} -2\ln\frac{\mathscr{L}(\mu,\hat{\hat{\boldsymbol{\theta}}})}{\mathscr{L}(\hat{\mu},\hat{\boldsymbol{\theta}})} & \hat{\mu}\geq 0 \\ -2\ln\frac{\mathscr{L}(\mu,\hat{\hat{\boldsymbol{\theta}}})}{\mathscr{L}(0,\hat{\hat{\boldsymbol{\theta}}})} & \hat{\mu}<0. \end{cases} \tag{4.29}$$

The test statistics used in the statistical interpretations of this thesis are detailed in Ref. [142], where the test of background-only hypothesis is stated as the problem of discover a positive signal, and the test of signal using CL_s technique is stated as the problem of setting upper limits. In the discovery of a positive signal, the test statistic is denoted as $q_0 \equiv \tilde{t}_0$, i.e.,

$$q_0 = \begin{cases} -2\ln\lambda(0) & \hat{\mu}\geq 0 \\ 0 & \hat{\mu}<0 \end{cases} = \begin{cases} -2\ln\frac{\mathscr{L}(0,\hat{\hat{\boldsymbol{\theta}}})}{\mathscr{L}(\hat{\mu},\hat{\boldsymbol{\theta}})} & \hat{\mu}\geq 0 \\ 0 & \hat{\mu}<0, \end{cases} \tag{4.30}$$

while in the setting of upper limits, that is denoted as $\tilde{q}_\mu$, which is

$$\tilde{q}_\mu = \begin{cases} -2\ln\tilde{\lambda}_\mu & \hat{\mu}\leq\mu \\ 0 & \hat{\mu}>\mu \end{cases} = \begin{cases} -2\ln\frac{\mathscr{L}(\mu,\hat{\hat{\boldsymbol{\theta}}})}{\mathscr{L}(0,\hat{\hat{\boldsymbol{\theta}}})} & \hat{\mu}\leq 0 \\ -2\ln\frac{\mathscr{L}(\mu,\hat{\hat{\boldsymbol{\theta}}})}{\mathscr{L}(\hat{\mu},\hat{\boldsymbol{\theta}})} & 0<\hat{\mu}\leq\mu \\ 0 & \hat{\mu}>\mu. \end{cases} \tag{4.31}$$

The p-value of the test is determined by the p.d.f (equivalently, c.d.f.) of the test statistics, referred to as $f(t_\mu|\mu')$, which indicates the p.d.f of test statistic t_μ under the hypothesis $H' : \mu = \mu'$. Accordingly, the significance Z and the upper limit μ_0^{up} at 95% confidence-level (CL_s) can be derived following the procedure mentioned above. The p.d.f of the test statistics can be sampled by randomly generating pseudo experiments, which can be very computationally expensive. However, it can also be approximately described by simple analytical functions studied in Ref. [142], based on the asymptotic approximation theorems established by Wilks [170] and Wald [171]. The approximation can only be used when the approximation is good enough, which is guaranteed by

[21] Here we adapt to the notation in [142]

the binning algorithm in this analysis. Here we only show the generic p.d.f for t_μ. That is

$$f(t_\mu|\mu') = \frac{1}{2\sqrt{t_\mu}}\frac{1}{\sqrt{2\pi}}\left[\exp\left(-\frac{1}{2}(\sqrt{t_\mu}+\sqrt{\Lambda})^2\right)+\exp\left(-\frac{1}{2}(\sqrt{t_\mu}-\sqrt{\Lambda})^2\right)\right], \tag{4.32}$$

which is a non-central chi-square distribution with one degree of freedom, where Λ is the non-centrality parameter

$$\Lambda = \frac{(\mu-\mu')^2}{\sigma^2}, \tag{4.33}$$

where σ is the variance of the maximum likelihood estimator $\hat{\mu}$. The asymptotic p.d.f(c.d.f.) of Equation (4.30) and Equation (4.31) can be found in [142] (Chapter 3).

The statistical interpolations discussed above are related to the observed results. The expected results are evaluated on an imaginary sample called 'Asimov dataset' which is made artificially from the prediction. Hence, the parameter estimations performed on this dataset will obtain the true values of the parameters. Using the Asimov dataset produced with a specified μ, one can find the median significance of signal with strength of μ and claim it as the sensitivity of discovering a positive signal. In the case of setting limit, using the Asimov dataset produced with $\mu = 0$, the median and $\pm 1/2\sigma$ (from Gaussian confidence intervals) upper limits can be scanned, these can be interpreted as the expected upper limits and the variations of the expected upper limits.

4.6 Results

This section shows the final results of the $\tau_{\text{had}}\tau_{\text{had}}$ decay channel, after performing the likelihood fits (post-fit results). As comprehensively discussed in Section 4.5.1, the likelihood model includes the SR (of $\tau_{\text{had}}\tau_{\text{had}}$) and the $Z + hf$ CR. The signal strength μ is the parameter of interest and all other parameters are considered as nuisance parameters. The signal strength can be translated into physical cross sections by

$$\mu = \frac{\sigma_{\text{meas}}}{\sigma}, \tag{4.34}$$

where σ_{meas} is the cross section to be measured and σ is the cross section that normalises the signal template.

Although the SR events are the same in the searches for different non-resonant and resonant signals, various MVA score distributions are used in the construction of the likelihood functions. Therefore, likelihood fit is performed for each search.

Table 4.18 shows the expected number of events after the likelihood fit based on the non-resonant BDT score under the background-only hypothesis in the inclusive signal region and in the highest two BDT score bins, referred to as 1$^{\text{st}}$ and 2$^{\text{nd}}$ highest bins. These bins are expected to bring major contributes to the sensitivity. The post-fit number of events are very close in non-resonance and resonance searches. The results are interpreted into CL_s-based upper limits on the HH production cross section. The methodology is described in Section 4.5.2.

Table 4.18: Post-fit expected event yields and uncertainties of signal and background processes in the inclusive SR and in the 1st and 2nd highest BDT score bins. The likelihood fit is performed on non-resonant BDT score distribution in the SR and the $m_{\ell\ell}$ distribution in the $Z + hf$ CR under background-only hypothesis. Due to correlations of the uncertainties on the backgrounds, the uncertainties on each background process can be larger than the total background uncertainty.

Process	Inclusive SR	2nd Highest	1st Highest
$Z \to \tau\tau + hf$	1500 ± 160	5.9 ± 0.89	2.2 ± 0.41
True-τ_{had} $t\bar{t}$	3000 ± 170	2.3 ± 0.40	0.54 ± 0.11
Fake-τ_{had} $t\bar{t}$	2100 ± 200	2.5 ± 0.43	0.28 ± 0.067
Multi-jet	1300 ± 137	1.5 ± 0.40	0.48 ± 0.17
Single Higgs boson	65 ± 12	2.6 ± 0.65	1.7 ± 0.45
Single top-quark	230 ± 68	0.36 ± 0.29	0.34 ± 0.20
Other backgrounds	200 ± 34	1.2 ± 0.31	0.43 ± 0.11
Total background	8400 ± 90	16 ± 1.6	6.0 ± 0.8
Data	8400	15	8
SM ggF HH	5.0 ± 0.84	0.87 ± 0.15	1.5 ± 0.25
SM VBF HH	0.15 ± 0.012	0.017 ± 0.0014	0.021 ± 0.0016

The difference between the searches for individual signals are embodied in the MVA models that are used to optimise in particular for that signal process. The distributions of the MVA input variables and the exemplary MVA output scores are illustrated in Figure 4.26 and Figure 4.27-4.29, including the BDT score for the non-resonant HH search and PNN evaluated at various mass points for the resonant $X \to HH$ search, where the event yields and uncertainties are obtained from the corresponding likelihood fits under background-only hypothesis. The distribution of data is compatible with the prediction of background-only hypothesis, except for the highest PNN score bins in the search for resonant $X \to HH$ with $m_X = 900$ and 1000 GeV. For these two mass hypothesis, the distributions of the PNN input variables of the events in the highest PNN score bins are shown in Figure 4.30 and 4.31 to visualise the excess.

The impact of different sources of uncertainties is evaluated by likelihood fits to data under signal + background hypothesis. The relative contribution of each source is shown in Table 4.19, for non-resonant HH and resonant $X \to HH$ at various mass points. The statistical uncertainty of data dominates the uncertainty. Theoretical uncertainty in single Higgs boson and SM HH are more important in the non-resonant HH case. Experimental and fake-τ_{had} background estimation uncertainties impact more on the lower mass points of resonant $X \to HH$.

4.6.1 Results in the search for non-resonant HH production

In the non-resonant search, the observed (expected) upper limit on the signal strength of SM HH at 95% confidence-level (CL) is 5.0 (4.4). As the baseline signal is the SM HH production via

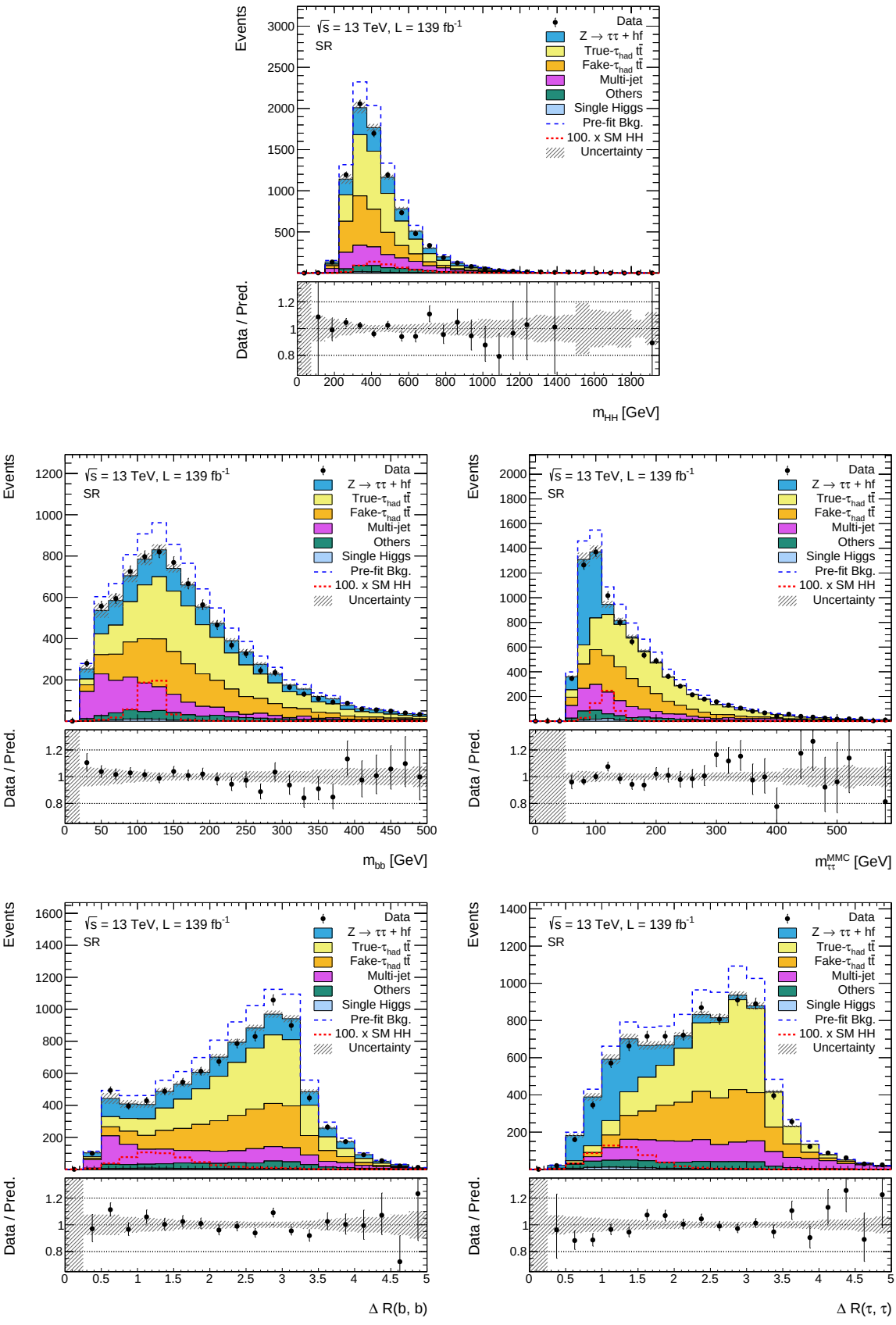


Figure 4.26: Distributions of input variables for MVA after the background-only-hypothesis likelihood fit. The non-resonant ggF+VBF *HH* signal is scaled by a factor of 100 with respected to its cross section. In the lower panel, the ratio of data and the total background prediction is displayed in black dots. The hatched area indicates the total systematic uncertainty of the backgrounds.

Table 4.19: Breakdown of the relative contribution of different sources of uncertainties in the best-fit signal strength determined by the likelihood fit, for non-resonant HH and resonant $X \to HH$ at various mass points. The percentage contributions are obtained by letting out a particular group of nuisance parameters in the likelihood fit and subtracting the obtained uncertainty in the best-fit signal strength in quadrature from the total uncertainty, and then dividing the difference by the total uncertainty. The sum in quadrature of the individual components differs from the total uncertainty due to correlations between the groups of uncertainties. The best-fit value and uncertainties of the signal strength is shown in the last row, the resonant $X \to HH$ events are normalised with respect to a cross section of 1 pb.

Source of uncertainty	Non-resonant HH	Resonant $X \to HH$		
		300 GeV	500 GeV	1000 GeV
Data statistical	87%	81%	94%	93%
Systematic	50%	58%	35%	36%
Z + HF normalisation	3%	4%	6%	2%
$t\bar{t}$ normalisation	4%	4%	6%	5%
MC statistical	24%	36%	25%	13%
Experimental	14%	25%	10%	16%
Jet/$E_{\mathrm{T}}^{\mathrm{miss}}$	8%	20%	7%	1%
b-tagging	5%	10%	3%	7%
τ_{had}	6%	8%	6%	13%
Others	3%	4%	3%	5%
Background estimation	8%	17%	4%	6%
Multi-jet	4%	4%	2%	1%
Fake-τ_{had} $t\bar{t}$	7%	16%	3%	6%
Theoretical	38%	28%	19%	29%
$Z + hf$	7%	15%	9%	12%
$t\bar{t}$	5%	17%	4%	4%
Single top-quark	7%	2%	5%	2%
Single Higgs boson	24%	17%	14%	25%
Other backgrounds	5%	4%	5%	7%
Signal	27%	17%	9%	22%
Best-fit $\mu \pm \sigma_\mu$	0.71 ±1.8	0.12 ±0.38	0.0064 ±0.021	0.023 ±0.011

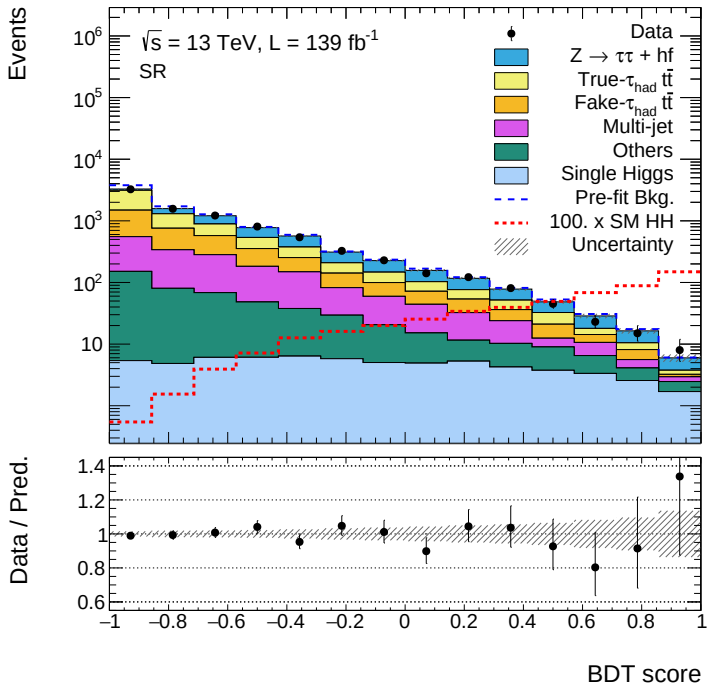


Figure 4.27: Distributions of the BDT score for the non-resonant HH search after the background-only-hypothesis likelihood fit. The signal is scaled by a factor of 100 with respected to the SM cross section. In the lower panel, the ratio of data and the total background prediction is displayed in black dots. The hatched area indicates the total systematic uncertainty of the backgrounds.

ggF and VBF, the observed (expected) upper limit on the SM HH production cross section [22] at 95% CL is 150 fb (130 fb). Together with the expected upper limit under $\pm 1, 2\text{-}\sigma$ variations, the numbers are summarised in Table 4.20. The observed upper limits are found to be compatible with the expected within 1σ variation.

Table 4.20: Observed and expected upper limits on the SM HH signal strength μ and production cross section σ_{HH}^{SM}. The expected limits under $\pm 1, 2\text{-}\sigma$ variations are also listed.

	Observed	-2σ	-1σ	Expected	1σ	2σ
95% upper limits on μ	5.0	2.4	3.2	4.4	6.2	8.2
95% upper limits on σ_{HH}^{SM} [fb]	150	70	95	130	180	240

4.6.2 Results in the search for resonant $X \to HH$ production

In the search for resonant $X \to HH$ production, upper limits are set on the cross section of $gg \to X \to HH$ at 95% CL for each mass hypothesis of the narrow-width scalar particle. The observed and expected upper limits, as well as the $\pm 1, 2\text{-}\sigma$ variations on the expected limits are visualised as a function of m_X in Figure 4.32, where the values are summarised in Table 4.21. Just as the

[22] To obtain the upper limit on the cross section, the uncertainties in the theoretical calculation of SM HH cross section, by which the number of SM HH signal events are normalised, is not considered.

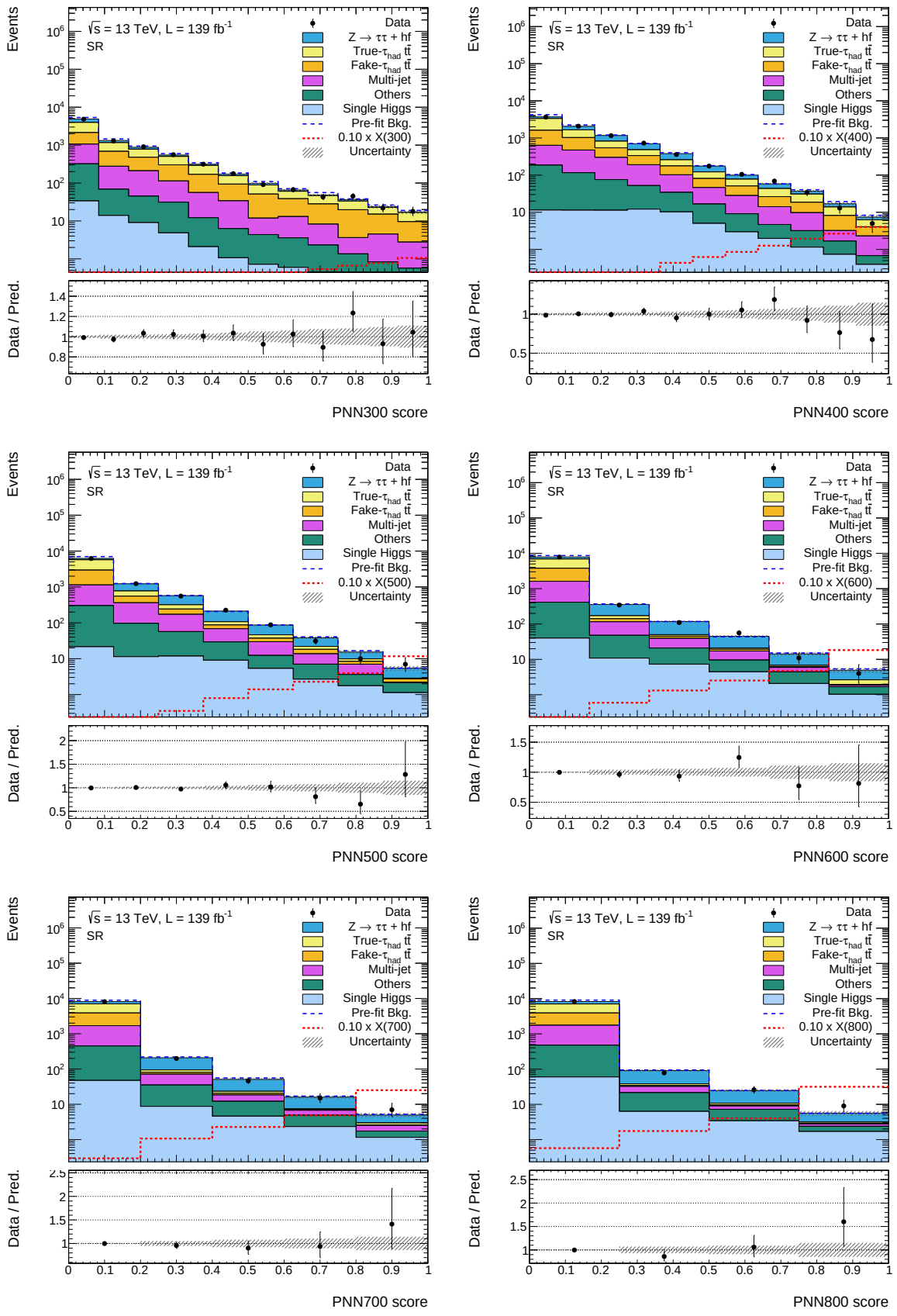


Figure 4.28: Distributions of the PNN evaluated at m_X = 300-800 GeV for the resonant $X \to HH$ search after the corresponding background-only-hypothesis likelihood fits. The cross section of the resonant HH signal is assumed to be 1 pb, and the signals are scaled by a factor of 0.1. In the lower panel, the ratio of data and the total background prediction is displayed in black dots. The hatched area indicates the total systematic uncertainty of the backgrounds.

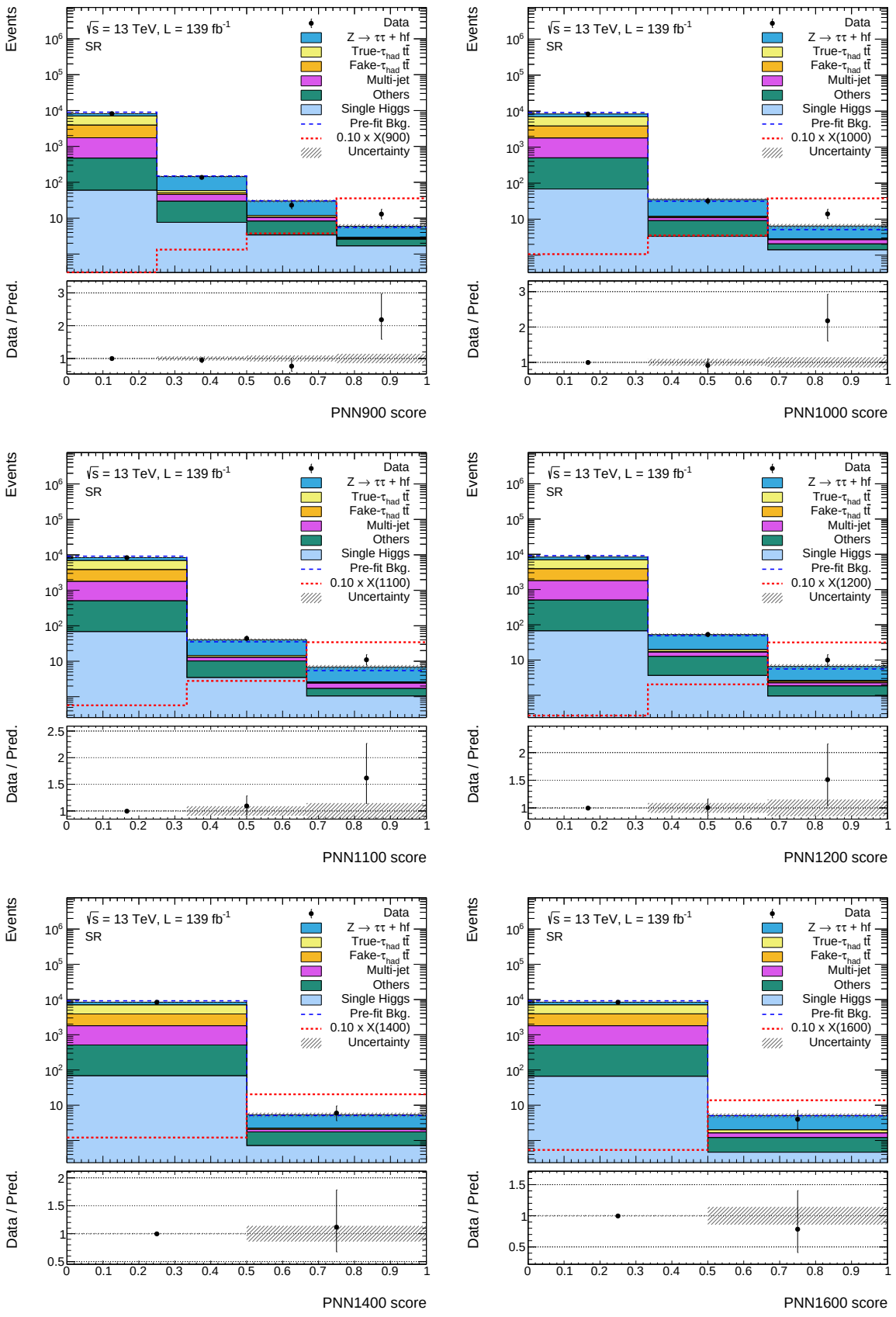


Figure 4.29: Distributions of the PNN evaluated at m_X = 900-1600 GeV for the resonant $X \to HH$ search after the corresponding background-only-hypothesis likelihood fits. The cross section of the resonant HH signal is assumed to be 1 pb, and the signals are scaled by a factor of 0.1. In the lower panel, the ratio of data and the total background prediction is displayed in black dots. The hatched area indicates the total systematic uncertainty of the backgrounds.

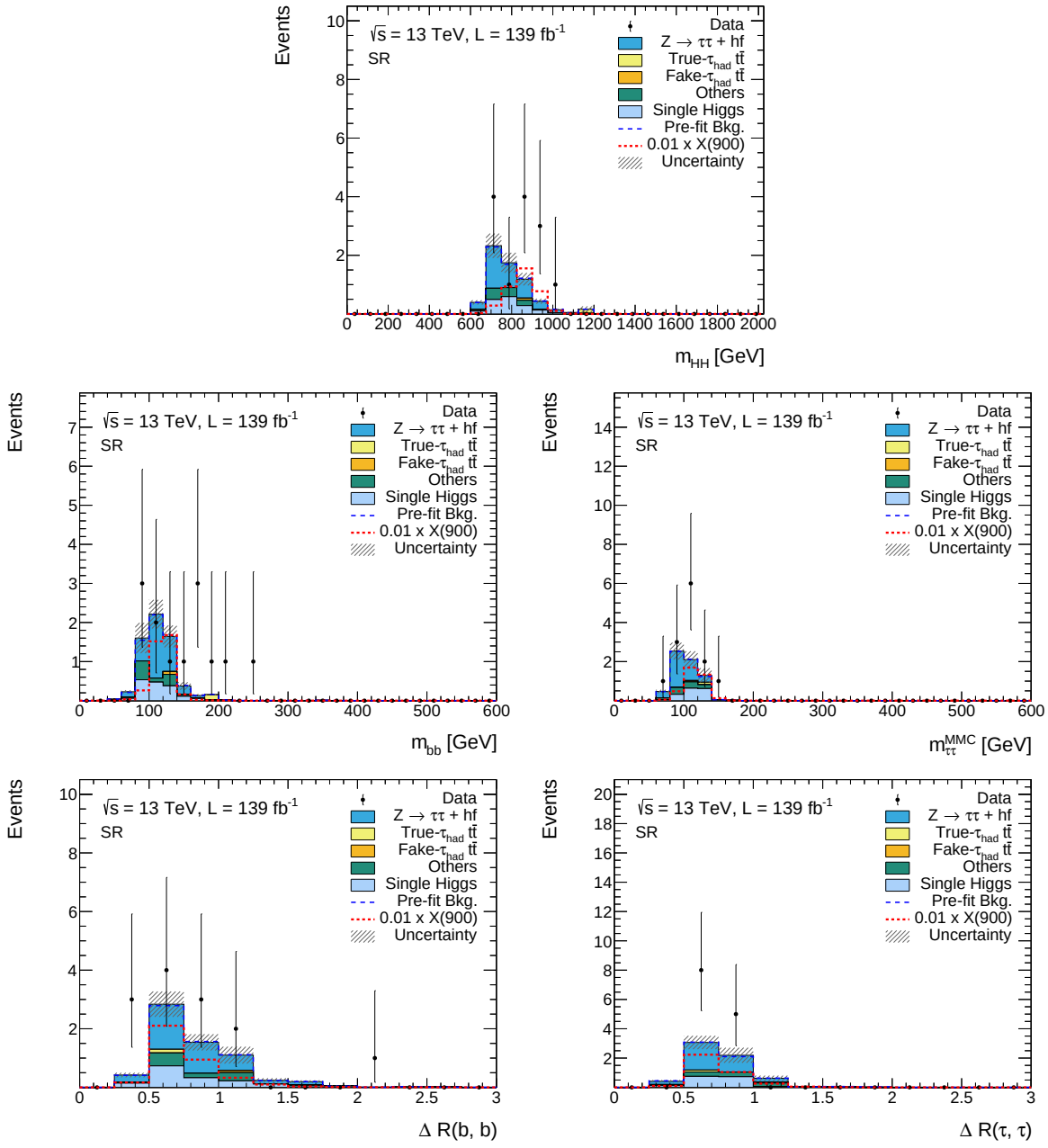


Figure 4.30: Distributions of input variables for MVA after the background-only-hypothesis likelihood fit, in the highest PNN(m_X=900 GeV) bin. The cross section of the resonant HH signal is assumed to be 1 pb, and the signals are scaled by a factor of 0.01. In the lower panel, the ratio of data and the total background prediction is displayed in black dots. The hatched area indicates the total systematic uncertainty of the backgrounds.

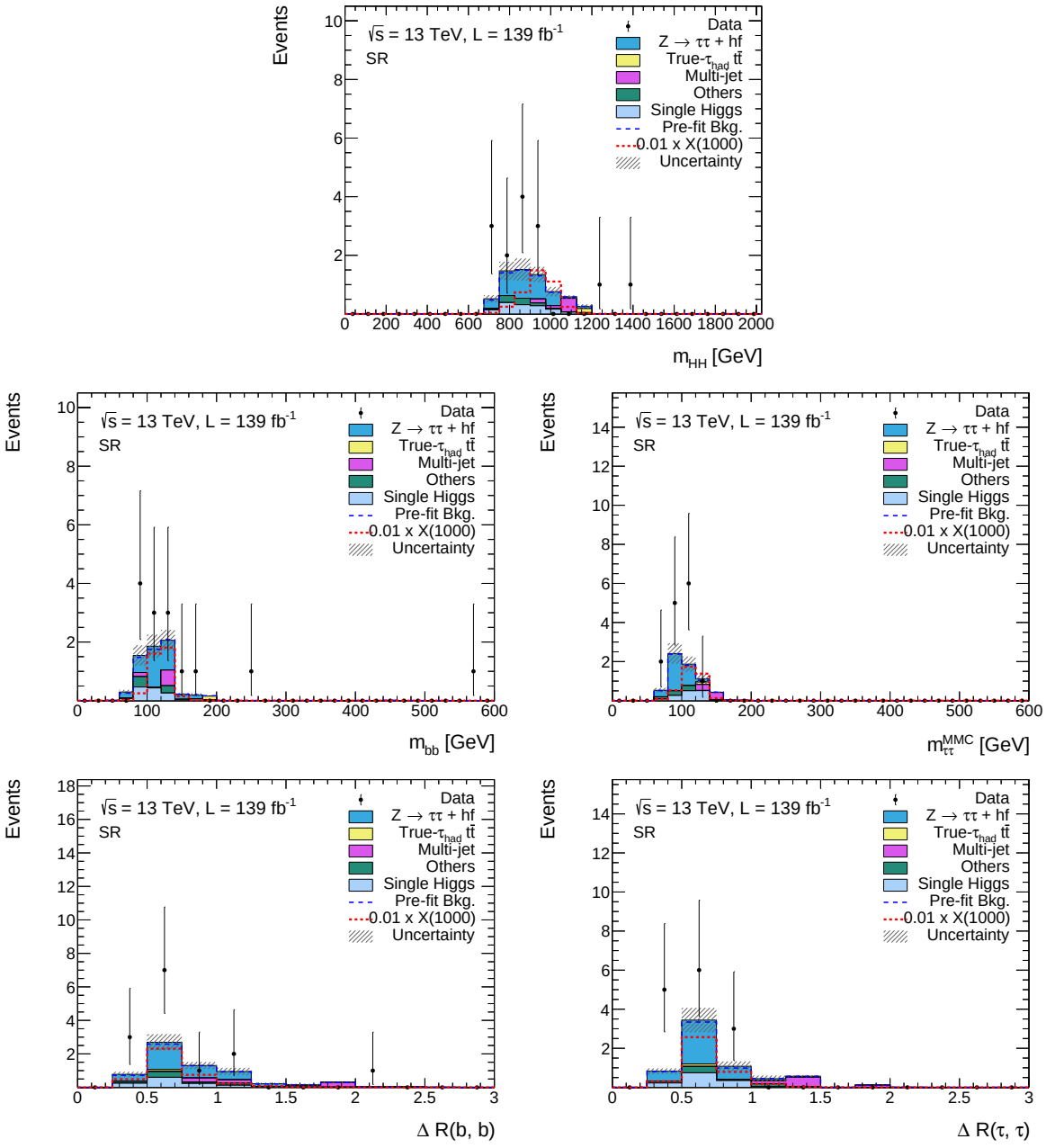


Figure 4.31: Distributions of input variables for MVA after the background-only-hypothesis likelihood fit, in the highest PNN(m_X=1000 GeV) bin. The cross section of the resonant HH signal is assumed to be 1 pb, and the signals are scaled by a factor of 0.01. In the lower panel, the ratio of data and the total background prediction is displayed in black dots. The hatched area indicates the total systematic uncertainty of the backgrounds.

patterns in the comparison of the MVA output distributions of data and background prediction, the observed upper limits are compatible with the expected upper limits for the searches of all mass points except for 900 GeV and 1000 GeV, where the observed limits are larger than the $+2\sigma$ expected limits. In fact, a wide excess of upper limits on the $X \to HH$ production is observed at m_X from 700 GeV to 1200 GeV.

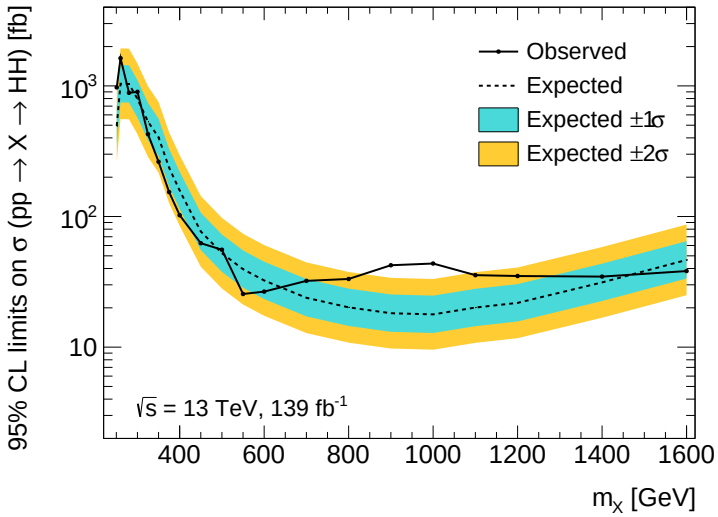


Figure 4.32: Observed (solid line) and expected (dashed line) upper limits on the cross section of $gg \to X \to HH$ at 95% CL in the search for resonant Higgs boson pair production via a narrow-width scalar particle (X) as a function of m_X. The markers indicate that the upper limits are calculated based on the signal templates generated by simulated samples with the corresponding mass. The intermediate upper limits are linearly interpolated. The green and yellow bands stand for the $\pm 1, 2$-σ variations on the expected limits, respectively.

4.6.3 Checks on the excess

This wide excess is expected if the $X \to HH$ signal with m_X equal 900 GeV or 1000 GeV exists. This is checked in Figure 4.33, where the limits are calculated by replacing the real dataset with the Asimov dataset build under the signal+background (S+B) prediction with the cross section of the signal set to the corresponding expected upper limits. The pattern originates from the fact that the same events are used for the statistical interpretation of the searches for all mass points. This approach brings correlated fluctuations of events between different mass points. The correlation effect is stronger for the events in the high PNN score region, where the number of events is lower and the same events are likely to be selected by the PNNs evaluated with nearby masses, especially for the high mass PNNs. As demonstrated in Figure 4.34, the PNN500 selects events with m_{HH} close to 500 GeV, while PNN1000 treats all events with high m_{HH} as signal-like events.

The results on the search for resonant $X \to HH$ production are also interpreted into the significance of discovering a new signal. Local and global significances are derived following the procedure introduced in Section 4.5.2. Figure 4.35 depicts the local p-value of the background-

Table 4.21: Summary of the observed and expected upper limits, as well as the expected upper limits under $\pm 1, 2\text{-}\sigma$ variations on the cross section of $gg \to X \to HH$ at 95% CL in the search for resonant Higgs boson pair production via a narrow-width scalar particle (X) of various m_X.

m_X [GeV]	95% CL upper limits on $\sigma_{gg\to X\to HH}$ [fb]					
	Observed	-2σ	-1σ	Expected	1σ	2σ
251	970	270	360	490	690	920
260	1600	560	750	1000	1500	1900
280	890	560	750	1000	1400	1900
300	900	430	570	800	1100	1500
325	430	290	380	530	740	990
350	260	220	290	400	560	750
375	150	130	170	240	330	440
400	100	85	110	160	220	300
450	62	41	55	77	107	140
500	56	28	38	52	73	98
550	26	21	28	40	55	74
600	27	17	23	32	45	60
700	32	13	17	24	33	45
800	33	11	15	20	28	38
900	42	9.8	13	18	25	34
1000	44	9.6	13	18	25	33
1100	36	11	14	20	28	37
1200	35	12	16	22	30	41
1400	35	17	23	31	44	58
1600	38	25	34	47	65	87

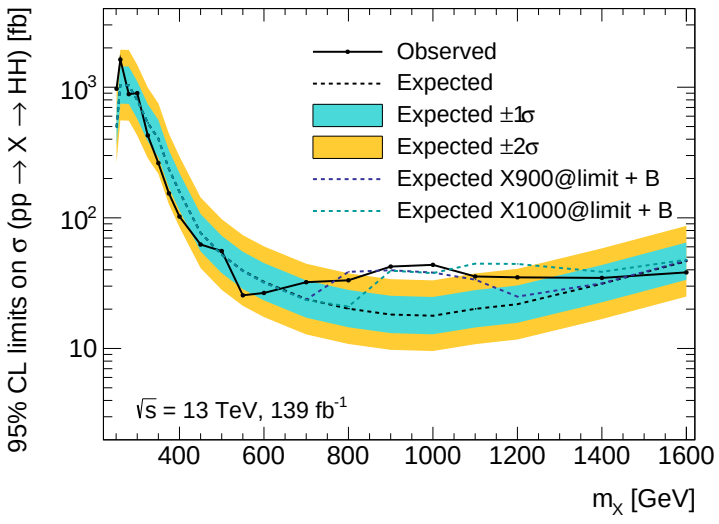
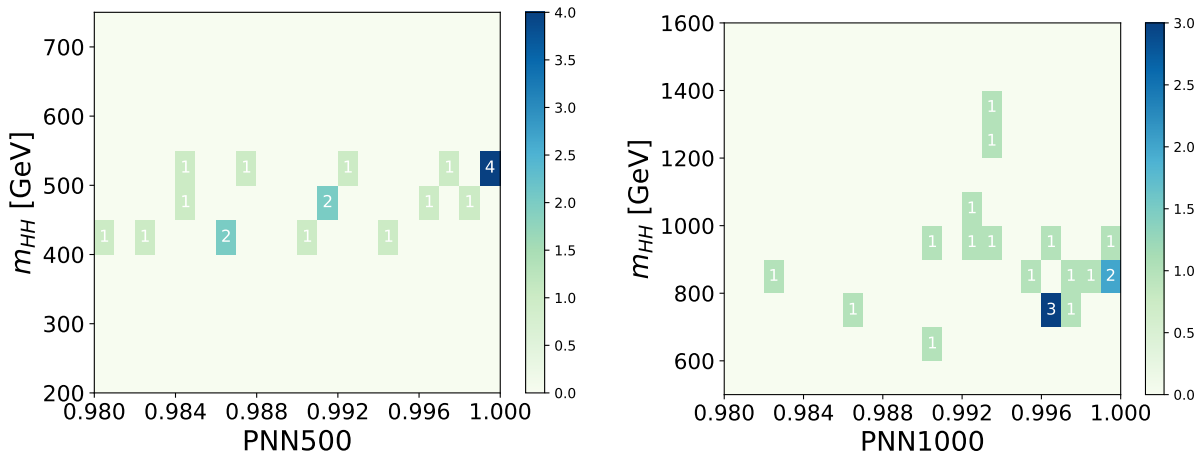


Figure 4.33: Observed (black solid) and expected (black dashed) upper limits on the cross section of $gg \to X \to HH$ at 95% CL in the search for resonant Higgs boson pair production via a narrow-width scalar particle (X) as a function of m_X. And the expected upper limits if $X \to HH$ signal with m_X equal 900 GeV (purple dashed) or 1000 GeV (cyan dashed) exists. The resonant signals are assumed to have cross sections equal to the upper limits.



(a) PNN500 (b) PNN1000

Figure 4.34: Two-dimensional distribution of PNN score and m_{HH} on data sample. The high PNN500 (a) and PNN1000 (b) score region are shown. The colour and number indicate the number of events.

only hypothesis, as a function of the mass of the scalar particle X, where the 1–4σ significances are indicated in orange horizontal lines. The most significance excess is at the mass hypotheses of $m_X = 1000$ GeV, with a local significance of 2.8σ and a global significance of 1.7 ± 0.02.

In the derivation of the global significance, the distribution of $Z_{\text{max}}^{\text{local}}$ is sampled by generating 10^4 pseudo experiments. Firstly, one needs to generate pseudo-data in the SR of all mass hypotheses. The pseudo-data follows the distribution predicted by the background-only hypothesis. The goal is to find the joint distribution of the counting experiments of all mass hypotheses (PNN score bins) while preserving the correlation between different counting experiments. Based on the Sklar's theorem [172], the joint distribution can be expressed in terms of its marginals and a *copula* C. The copula is a multivariate cumulative distribution function (c.d.f), in which each variable $(U_1, U_2, \ldots, U_d)$ has a uniform marginal p.d.f on $[0, 1]$. It can be written as

$$C(u) = Pr[U_1 \leq u_1, U_2 \leq u_2, \ldots, U_d \leq u_d]. \tag{4.35}$$

The copula is a useful tool to describe the correlation between random variables. Pseudo-data can be generated given a certain copula [173]. A Gaussian copula is used in this analysis. It is

$$C_R^{Gauss}(u) = \Phi_R(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_d)), \tag{4.36}$$

where R is the input correlation matrix $R \in [-1, 1]^{d \times d}$, the Φ_R is the c.d.f of the multivariate normal distribution with covariance matrix equals to the correlation matrix R, and Φ^{-1} is the inverse c.d.f of normal distribution. The correlation matrix R can be estimated based on the expected backgrounds. For Poisson counting experiments i and j with parameters λ_i and λ_j set to the post-fit [23] expected

[23] Likelihood fit based on non-resonant HH BDT and background-only hypothesis.

number of background event, and random variables X_i and X_j, the correlation coefficient of X_i and X_j can be derived as

$$\rho_{ij} = \frac{\lambda_{i\cap j}}{\sqrt{\lambda_i \lambda_j}}, \tag{4.37}$$

where $\lambda_{i\cap j}$ is the parameter of the Poisson counting experiment for the overlapped events between i and j. The correlation coefficient between any pair of mass hypotheses and PNN score bins can be calculated, the covariance matrix R is thus formed, part of R is shown in Figure B.10. Once the Gaussian copula is constructed, a vector uniformly distributed random variables $(U_1, U_2, \ldots, U_d)$ can be generated. Then the pseudo-data can be constructed by

$$(X_1, X_2, \ldots, X_d) = \left(F_1^{-1}(U_1), F_2^{-1}(U_2), \ldots, F_d^{-1}(U_d)\right), \tag{4.38}$$

where F_i^{-1} are the inverse c.d.f of the desired p.d.f, which is Poisson counting p.d.f of that bin. For the ZCR, a bootstrap method [174] is used to generate the pseudo-data (see Appendix B.4). The observables correspond to the auxiliary experiments are also resampled randomly according to their constrain terms, and the fluctuations of these observables are correlated between different mass hypotheses by the bootstrap method. The same background and signal templates are used to calculate the local significances for each pseudo-experiment. Finally, the distribution of $Z_{\text{max}}^{\text{local}}$ is sampled and $Z^{\text{glob.}}$ is derived by Equation (4.24), as shown in Figure 4.36. The quoted uncertainty of $Z^{\text{glob.}}$ is due to the finite number of pseudo-experiments.

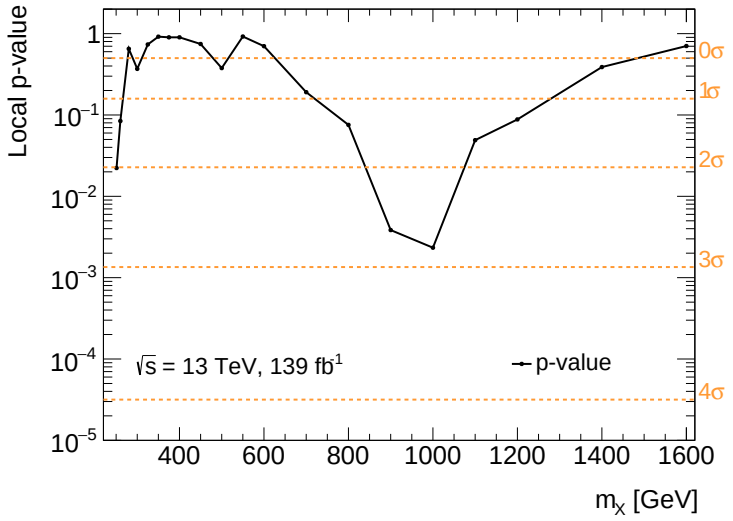


Figure 4.35: P-value of the background-only hypothesis in the search for resonant Higgs boson pair production via a narrow-width scalar particle (X) as a function of m_X. The markers indicate that the p-values derived based on the signal templates generated by simulated samples with the corresponding mass. The intermediate p-values are linearly interpolated. The dashed orange lines stand for the converted significance of local excess up to 4σ.

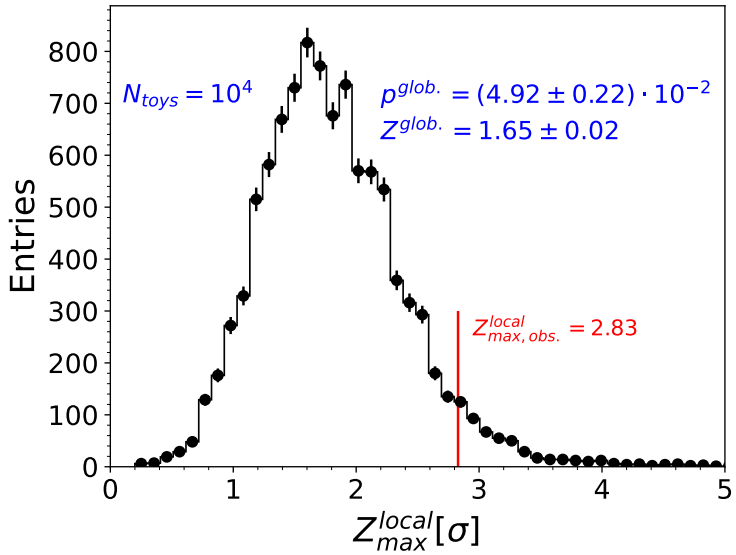


Figure 4.36: Distribution of the maximum local significance of 10^4 pseudo-experiments. The observed maximum local significance (at $m_X = 1000$ GeV mass point) is indicated in the red vertical line. The corresponding global p-value and significance are shown in blue text. The vertical error bar shows the statistical uncertainty of the pseudo-data.

4.7 Combination of $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels

4.7.1 Overview of $\tau_{\text{lep}}\tau_{\text{had}}$ channel

The results of the full $b\bar{b}\tau^+\tau^-$ decay channel are given by combining the results of the $\tau_{\text{had}}\tau_{\text{had}}$ and the $\tau_{\text{lep}}\tau_{\text{had}}$ decay channels. The $\tau_{\text{lep}}\tau_{\text{had}}$ channel is very similar as the $\tau_{\text{had}}\tau_{\text{had}}$ channel, but with one of the τ-lepton decaying leptonically. The analysis strategy of this decay channel is outlined below.

The $\tau_{\text{lep}}\tau_{\text{had}}$ events are triggered by single-lepton triggers (SLT) or lepton+τ_{had} triggers (LTT). In $\tau_{\text{lep}}\tau_{\text{had}}$ channel, the events are split into these two trigger categories. The criteria to select signal candidates are identical to those in $\tau_{\text{had}}\tau_{\text{had}}$ channel, but exact one electron or muon and one τ_{had} is required instead of two τ_{had}. Furthermore, m_{bb} is required to be less than 150 GeV to reject the contamination of $t\bar{t}$ background. The $\tau_{\text{lep}}\tau_{\text{had}}$ SR selections are as summarises as follows:

- Pass single lepton trigger (denote as SLT, same SET, SMT) or lepton + τ_{had} trigger (denote as LTT, including electron (ETT) and muon (MTT) + τ_{had} triggers; For SLT the p_{T} thresholds are identical to those used in Table 4.6 for $Z+hf$ CR; For LTT the p_{T} thresholds are summarised in Table 4.22;
- Same jet selection as in $\tau_{\text{had}}\tau_{\text{had}}$. LTT trigger has analogous additional jet requirements as DTT;
- Exactly one *Tight* electron or one *Medium* muon, plus one *Loose* τ_{had};
- Same event-level selections as $\tau_{\text{had}}\tau_{\text{had}}$ (e/μ and τ_{had} have opposite-sign charges);
- Additional requirement of $m_{bb} < 150$ GeV to reject $t\bar{t}$ background.

Table 4.22: The p_{T} thresholds of the lepton + τ_{had} triggers in $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel. The online and offline p_{T} thresholds are displayed outside and inside the brackets. The specific configurations of the triggers depend on the data-taking period. The form 'A + B' indicates $p_{\mathrm{T}}^{\tau_{\mathrm{had}}}$ + p_{T}^{ℓ} ($\ell = e/\mu$).

Trigger	Online (Offline) p_{T} threshold [GeV]
ETT	25 (30) + 17 (18)
MTT	25 (30) + 14 (15)

Slightly different binning optimisation algorithms and MVA algorithms are used in this channel, but the purposes are the same. In the search for non-resonant HH, a deep neural network (NN) is used to extract the signals, based on input features constructed by the momenta of the final state objects. Similar composition of backgrounds as the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel is expected, however, the fraction of contribution from the $t\bar{t}$ background is significantly larger in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel. The multi-jet and $t\bar{t}$ events with τ_{had} misidentified by quark/gluon-initiated jets are estimated by a data-driven method which is similar as the fake factor method discussed in Section 4.3.3. The $t\bar{t}$ *FF*s and the multi-jet *FF*s are calculated separately. The reweighting of true-$t\bar{t}$ described in Section 4.3.2.1 is applied in the derivation of the *FF*s. The *FF*s for fake-τ_{had}'s from $t\bar{t}$ and multi-jet are combined into one *FF* by a factor called r_{QCD}. Nuisance parameters arise from the uncertainties of this fake-τ_{had} estimation are summarised in Table 4.23. The contributions of the other background processes are estimated by simulated samples. The systematic uncertainties are derived following the same procedure as in the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel.

Table 4.23: Summary of systematic uncertainties related to fake-τ_{had} $t\bar{t}$ and multi-jet estimations.

Name of uncertainty	**Description**
FFStatTtbar	Statistical uncertainty of the $t\bar{t}$ *FF*s
FFStatQCD	Statistical uncertainty of the multi-jet *FF*s
Subtraction_bkg	Uncertainties in the subtracted background when applying the combined *FF*s
ttReweighting	Total uncertainty of the $t\bar{t}$ reweighting
FFStatrQCD	Statistical uncertainty of the r_{QCD}
FFVarrQCD	Conservative uncertainty that considers the possible variation range (0–0.5) of the r_{QCD}

4.7.2 Treatment of $t\bar{t}$ parameters in the combination

Eventually, the signal regions from the two trigger categories of the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ channel are integrated into the likelihood function Equation (4.13). The bins in each SR are added to the counting experiment and the parameters are correlated between two $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ SRs, the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ SR and the $Z + hf$ CR. The correlation of the normalisation factors and nuisance parameters of the $t\bar{t}$ background in different region needs to be treated with caution, because the kinematics of $t\bar{t}$ are diverse in these regions. This can be seen from the distribution of the scalar sum of the transverse momentums of the visible final state objects (S_T) as depicted in Figure 4.37. It shows that the S_{T} of $t\bar{t}$ in the $\tau_{\mathrm{lep}}\tau_{\mathrm{had}}$ SRs is softer than the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ and the $Z + hf$ CR. On the other hand, because a m_{bb} sideband is

selected in the $Z + hf$ CR, it may also cause differences in the kinematics of $t\bar{t}$. Therefore, three different $t\bar{t}$ normalisation correlation schemes are tested:

- Correlate all regions (scheme A);
- Decorrelate $\tau_{\text{lep}}\tau_{\text{had}}$ SRs from other regions (scheme B);
- Decorrelate $Z + hf$ CR from the signal regions (scheme C).

Likelihood fits are performed for each scheme, Table 4.24 summaries the best-fit $t\bar{t}$ normalisation factors. Despite the discrepancies in $t\bar{t}$ normalisation factors, the post-fit yields of $t\bar{t}$ under the three schemes agree with each other within the total uncertainty of $t\bar{t}$. Expected upper limits on non-resonant and resonant HH production cross sections are also derived based on each scheme, the results are also found to be compatible with each other. The scheme A is chosen for clarity. Besides, each nuisance parameter of $t\bar{t}$ is checked by comparing their best-fit values and 1σ intervals between correlated and uncorrelated scenarios. The $t\bar{t}$ PS uncertainty shows the largest difference between $\tau_{\text{lep}}\tau_{\text{had}}$ and $\tau_{\text{had}}\tau_{\text{had}}$ SRs, as summarised in Table 4.25. It is more constrained in the $\tau_{\text{lep}}\tau_{\text{had}}$ SRs, especially in the SLT SR. To avoid unduly introducing constraints on the $t\bar{t}$ PS uncertainty, also considering the different phase spaces and contributions of $t\bar{t}$ in the $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ SRs, this nuisance parameter is uncorrelated between the three SRs.

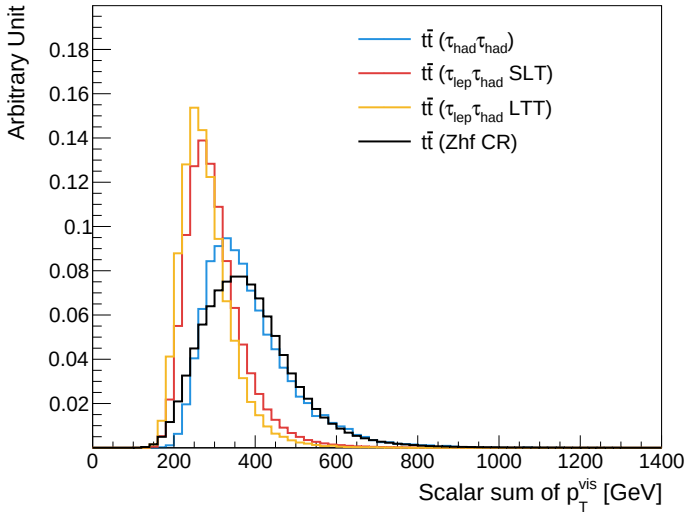


Figure 4.37: Distribution of the scalar sum of the transverse momentums of the visible final state objects in different regions.

4.7.3 Combined results

Other technical details about the combined likelihood fit can be found in Appendix B.5.1. Figure 4.38–4.41 shows the ranking of the nuisance parameters obtained in the non-resonant and several exemplary resonant likelihood fits to data. The nuisance parameters are ordered by the impact on the $\hat{\mu}$ (expressed as $\Delta\mu$), divided by the confidence interval of $\hat{\mu}$ (expressed as $\Delta\mu_{\text{tot}}$).

Table 4.24: Best-fit $t\bar{t}$ normalisation factors of each scheme. The likelihood fit is performed with non-resonant likelihood function under background-only hypothesis.

Scheme	Normalisation factor $\hat{\mu}_{t\bar{t}} \pm \Delta\mu_{t\bar{t}}$			
	$Z + hf$ CR	$\tau_{\text{had}}\tau_{\text{had}}$ SR	$\tau_{\text{lep}}\tau_{\text{had}}$ LTT SR	$\tau_{\text{lep}}\tau_{\text{had}}$ SLT SR
Scheme A	0.96 ± 0.04			
Scheme B	0.97 ± 0.04		0.92 ± 0.05	
Scheme C	0.99 ± 0.04	0.88 ± 0.04		

Table 4.25: Best-fit nuisance parameter related to $t\bar{t}$ PS uncertainties in $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ SRs. The 'Correlated' means the nuisance parameter is used publicly for all SRs. The likelihood fit is performed with non-resonant likelihood function under background-only hypothesis.

Correlated	$\tau_{\text{had}}\tau_{\text{had}}$ SR	$\tau_{\text{lep}}\tau_{\text{had}}$ LTT SR	$\tau_{\text{lep}}\tau_{\text{had}}$ SLT SR
-0.12 ± 0.30	-0.21 ± 0.88	-0.13 ± 0.62	-0.18 ± 0.33

Table 4.26 shows the expected number of events after the likelihood fit to the combined likelihood model based on the non-resonant BDT score ($\tau_{\text{had}}\tau_{\text{had}}$) and NN score ($\tau_{\text{lep}}\tau_{\text{had}}$) under the background-only hypothesis in each signal region, while Figure 4.42 shows the distributions of the MVA output scores after the likelihood fits to the corresponding BDT/NN/PNNs.

Table 4.26: Post-fit expected event yields and uncertainties of signal and background processes in the SRs of $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels. The likelihood fit is performed on non-resonant BDT score and NN score distribution in the $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ (SLT and LTT categories) SRs and the $m_{\ell\ell}$ distribution in the $Z + hf$ CR under background-only hypothesis. Due to correlations of the uncertainties on the backgrounds, the uncertainties on each background process can be larger than the total background uncertainty.

Process	$\tau_{\text{had}}\tau_{\text{had}}$ SR	$\tau_{\text{lep}}\tau_{\text{had}}$ SLT SR	$\tau_{\text{lep}}\tau_{\text{had}}$ LTT SR
$Z \to \tau\tau + hf$	1600 ±160	2600 ±190	580 ±61
True-τ_{had} $t\bar{t}$	3000 ±150	57000 ±1300	3800 ±190
Fake-τ_{had} $t\bar{t}$	2000 ±200	-	-
Multi-jet	1300 ±130	-	-
Fake-τ_{had} (Multi-jet and $t\bar{t}$)	-	34000 ±1500	1600 ±170
Single Higgs boson	66 ±13	150 ±18	23 ±4.3
Other backgrounds	460 ±76	4100 ±500	260 ±42
Total background	8400 ±90	98000 ±390	6400 ±79
Data	8400	98000	6400
SM ggF HH	5.0 ±0.83	5.7 ±0.99	1.4 ±0.24
SM VBF HH	0.15 ±0.012	0.19 ±0.016	0.053 ±0.0034

The same statistical interpretations as those used in the $\tau_{\text{had}}\tau_{\text{had}}$ channel are performed on the combined model. Upper limits are derived for the non-resonant HH production cross section in individual channels and in the combination, which are summarised in Table 4.27. The observed

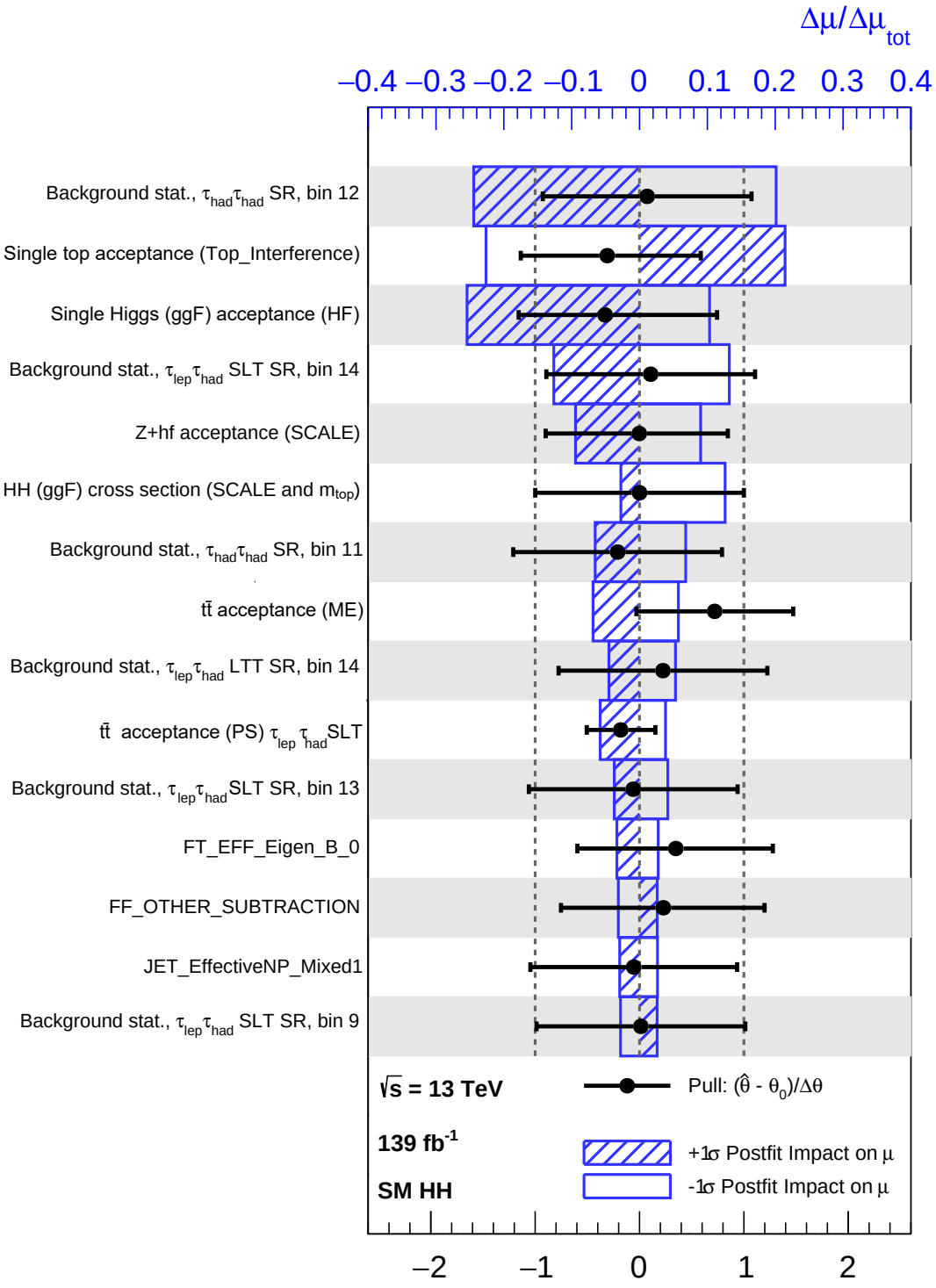


Figure 4.38: Ranking of the nuisance parameters obtained in the non-resonant likelihood fits to data, ordered by the impact on the $\hat{\mu}$ divided by the confidence interval of $\hat{\mu}$ ($\Delta\mu/\Delta\mu_{\text{tot}}$). The $\pm1\sigma$ impacts are expressed by hatched and and empty blue box, respectively. The black error bars shows the pulls of the nuisance parameters, which stands for the degree of deviation between the best-fit value and the value from the corresponding auxiliary measurements. The description of the nuisance parameters can be found in the text.

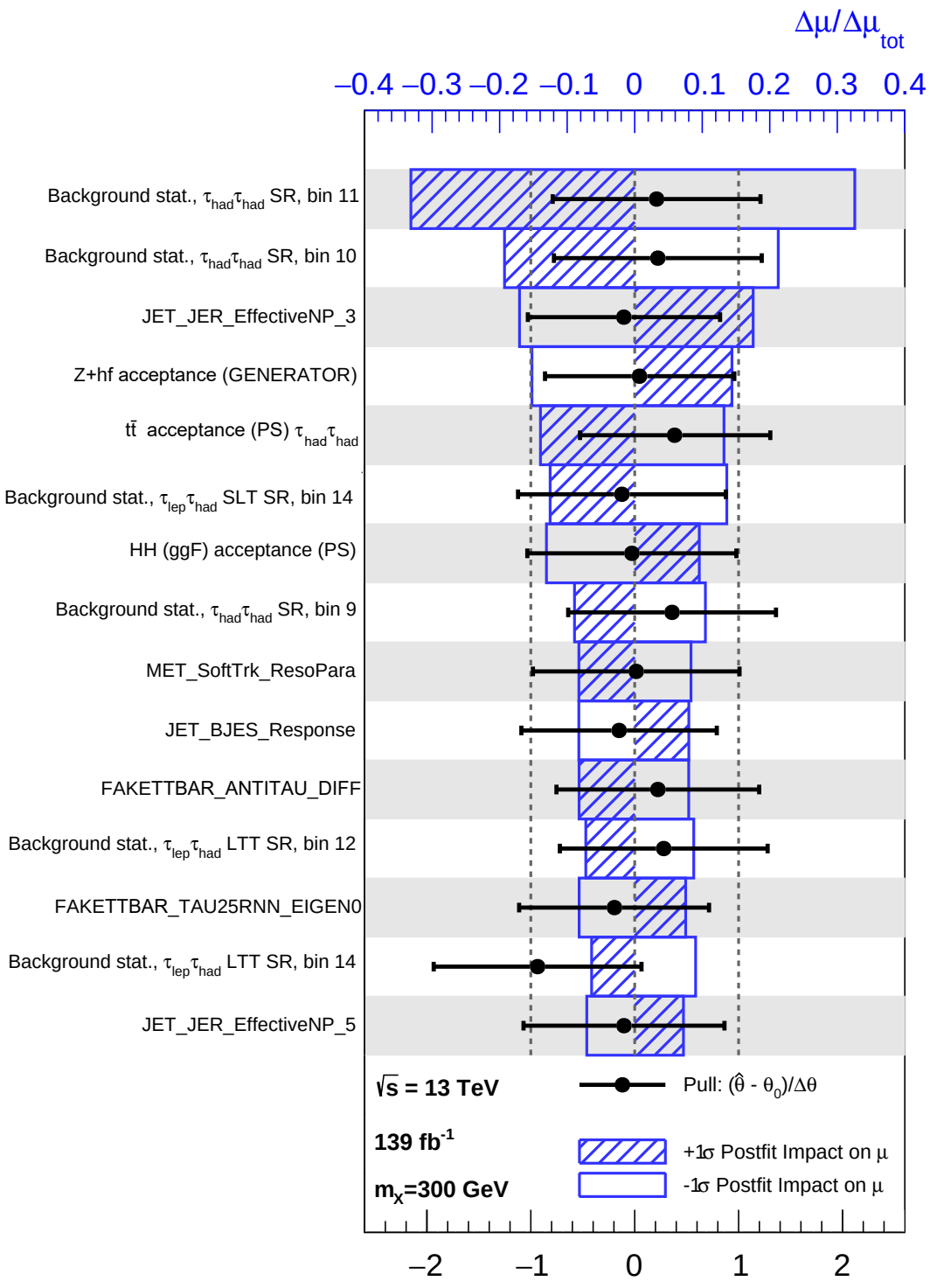


Figure 4.39: Ranking of the nuisance parameters obtained in the $m_X = 300$ GeV resonant likelihood fits to data, ordered by the impact on the $\hat{\mu}$ divided by the confidence interval of $\hat{\mu}$ ($\Delta\mu/\Delta\mu_{\text{tot}}$). The $\pm 1\sigma$ impacts are expressed by hatched and and empty blue box, respectively. The black error bars shows the pulls of the nuisance parameters, which stands for the degree of deviation between the best-fit value and the value from the corresponding auxiliary measurements. The description of the nuisance parameters can be found in the text.

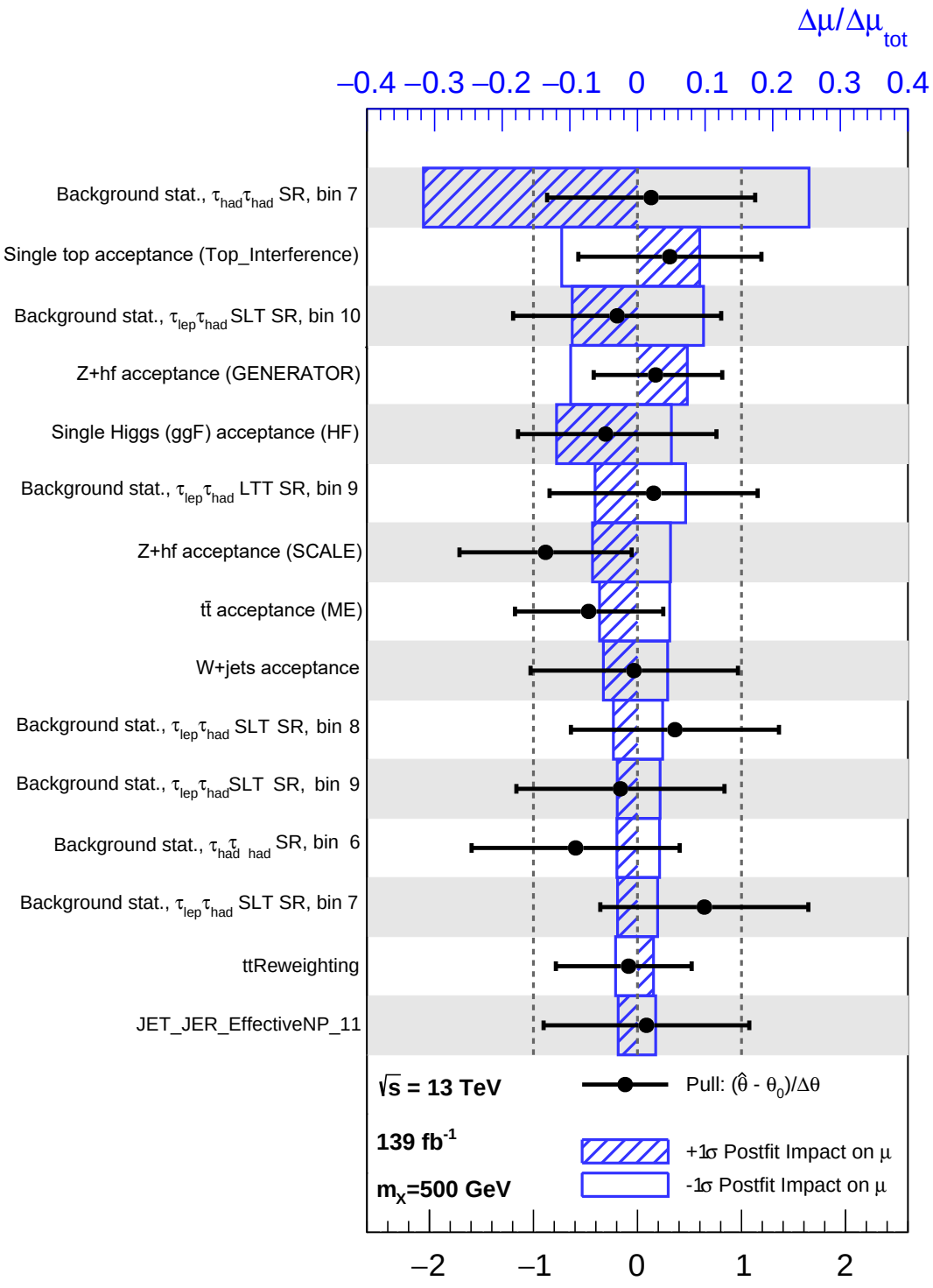


Figure 4.40: Ranking of the nuisance parameters obtained in the $m_X = 500$ GeV resonant likelihood fits to data, ordered by the impact on the $\hat{\mu}$ divided by the confidence interval of $\hat{\mu}$ ($\Delta\mu/\Delta\mu_{\text{tot}}$). The $\pm 1\sigma$ impacts are expressed by hatched and and empty blue box, respectively. The black error bars shows the pulls of the nuisance parameters, which stands for the degree of deviation between the best-fit value and the value from the corresponding auxiliary measurements. The description of the nuisance parameters can be found in the text.

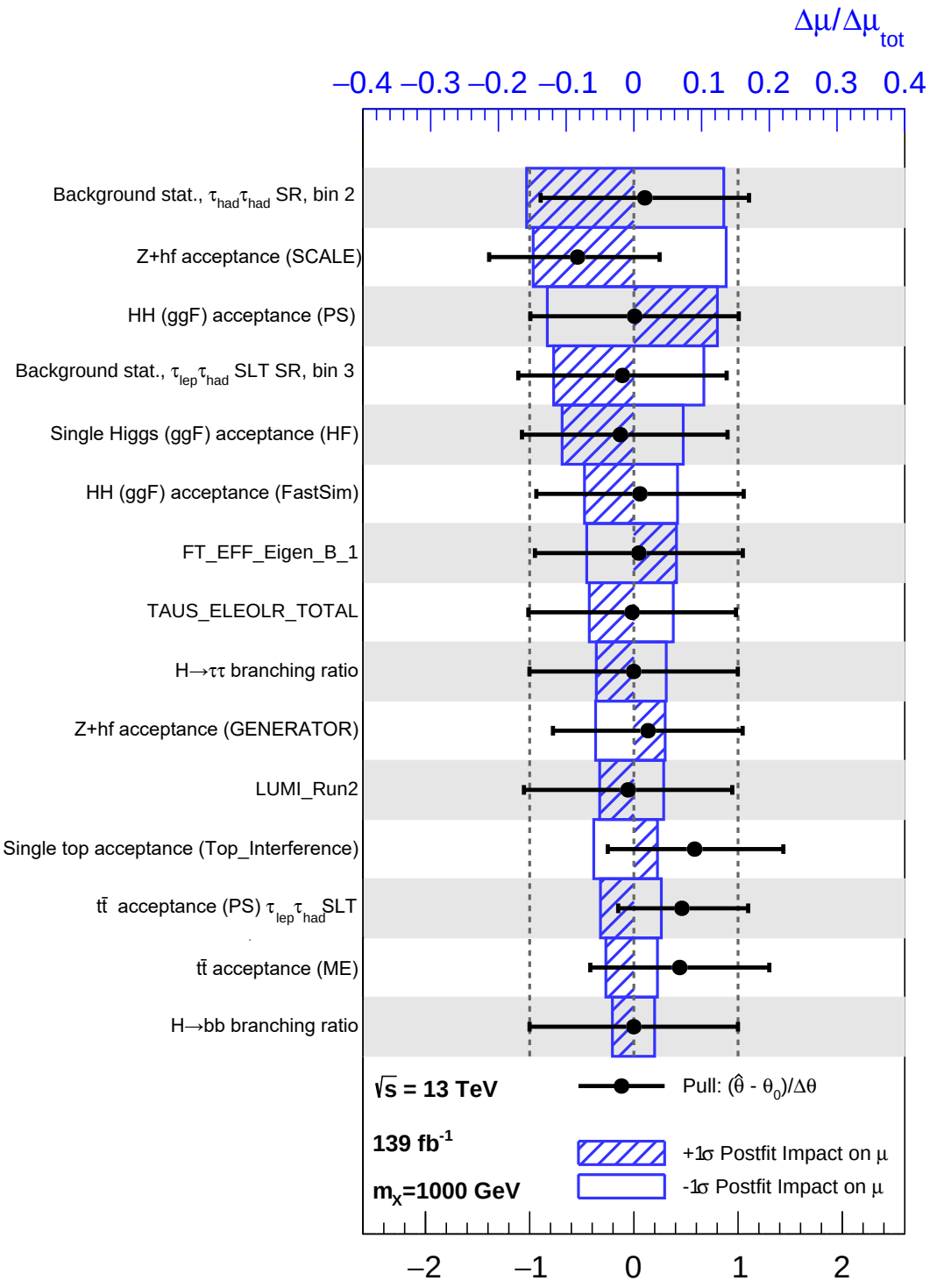


Figure 4.41: Ranking of the nuisance parameters obtained in the $m_X = 1000$ GeV resonant likelihood fits to data, ordered by the impact on the $\hat{\mu}$ divided by the confidence interval of $\hat{\mu}$ ($\Delta\mu/\Delta\mu_{\text{tot}}$). The $\pm 1\sigma$ impacts are expressed by hatched and and empty blue box, respectively. The black error bars shows the pulls of the nuisance parameters, which stands for the degree of deviation between the best-fit value and the value from the corresponding auxiliary measurements. The description of the nuisance parameters can be found in the text.

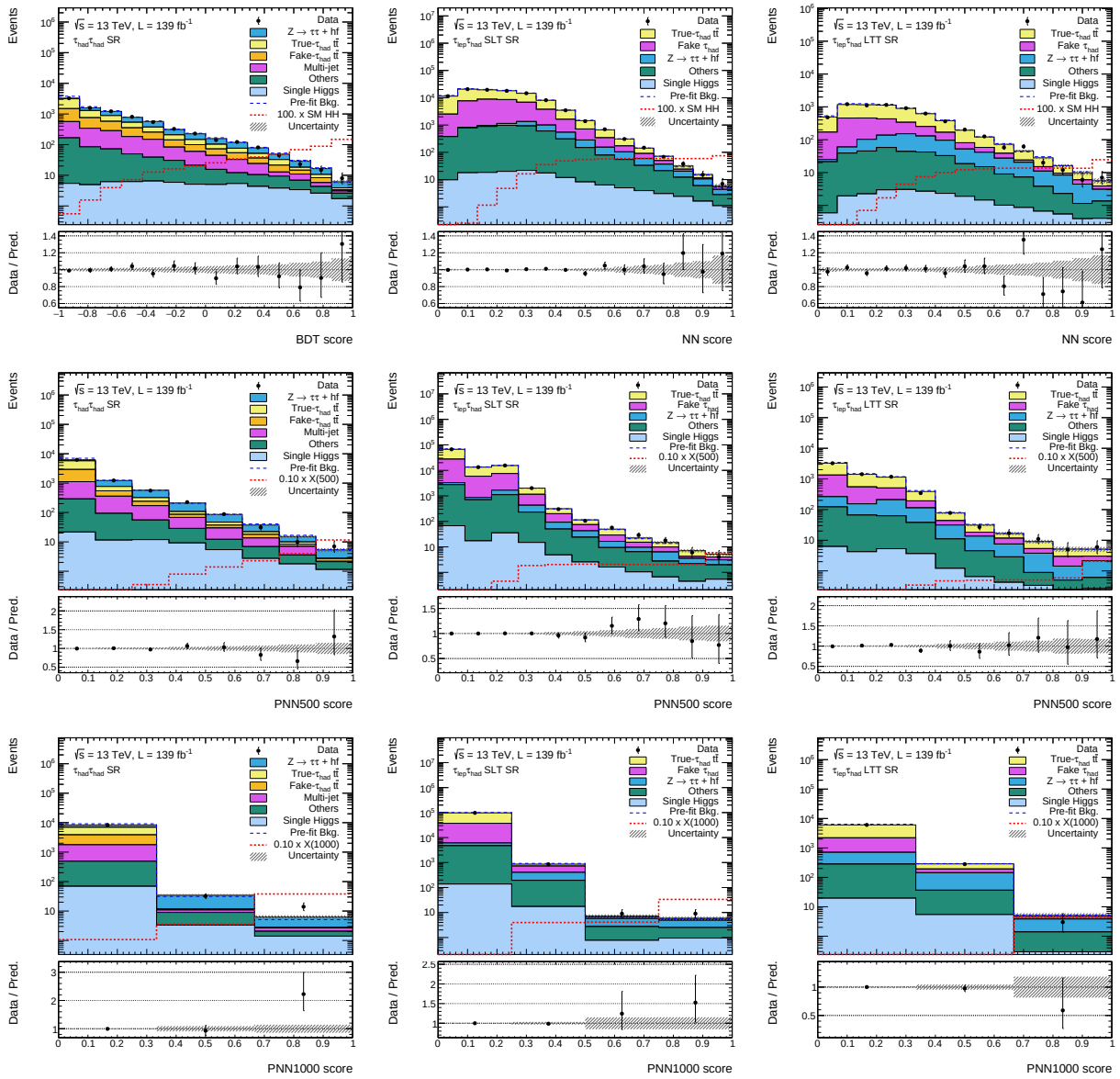


Figure 4.42: Distributions of the BDT/NN scores for the non-resonant HH search and PNN evaluated at $m_X = 500$ and 1000 GeV for the resonant $X \to HH$ search after the corresponding background-only-hypothesis likelihood fits in $\tau_{\text{had}}\tau_{\text{had}}$ SR (left), $\tau_{\text{lep}}\tau_{\text{had}}$ SLT (middle) SR and $\tau_{\text{lep}}\tau_{\text{had}}$ LTT SR (right). The signals are scaled to their expected upper limits. In the lower panel, the ratio of data and the total background prediction is displayed in black dots. The hatched area indicates the total systematic uncertainty of the backgrounds.

(expected) upper limit on the signal strength at 95% CL is 4.7 (3.9) and the observed (expected) upper limit on the SM HH production cross section is 140 fb (120 fb). The observed and expected are compatible within the 1σ variation.

For resonant $X \to HH$ production, upper limits and significances are obtained for individual channels and for the combination as well. The results are visualised in Figure 4.43 and in Figure 4.44 for the upper limits on the cross section of $gg \to X \to HH$ at 95% CL as a function of the mass of the scalar particle X and the local significances of observing the particle (local p-value of the background-only hypothesis), respectively. There is a wide excess at m_X from 700 to 1200 GeV driven by the results from $\tau_{\text{had}}\tau_{\text{had}}$ channel as discussed in Section 4.6. The most significance excess is also at the mass hypotheses of $m_X = 1000$ GeV, with a local significance of 3.0σ and a global significance of 2.0 ± 0.03. The global significance is derived using the same method as that in the $\tau_{\text{had}}\tau_{\text{had}}$ channel.

Table 4.27: Observed and expected upper limits on the SM HH signal strength μ in individual decay channels and in the combination. The expected limits under $\pm 1, 2\text{-}\sigma$ variations are also listed.

95% upper limits on μ	Observed	-2σ	-1σ	Expected	1σ	2σ
$\tau_{\text{had}}\tau_{\text{had}}$	5.0	2.4	3.2	4.4	6.2	8.2
$\tau_{\text{lep}}\tau_{\text{had}}$	9.6	4.2	5.6	7.8	11	15
Combination	4.7	2.1	2.8	3.9	5.4	7.2

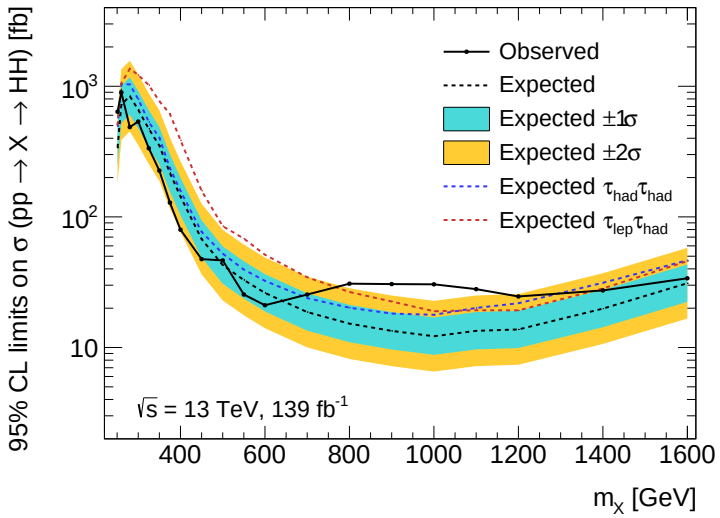


Figure 4.43: Observed (solid line) and expected (dashed line) upper limits on the cross section of $gg \to X \to HH$ at 95% CL in the search for resonant Higgs boson pair production via a narrow-width scalar particle (X) as a function of m_X. The markers indicate that the upper limits are calculated based on the signal templates generated by simulated samples with the corresponding mass. The intermediate upper limits are linearly interpolated. The black, blue and red lines corresponds to the results in the combination, the standalone $\tau_{\text{had}}\tau_{\text{had}}$ channel and the standalone $\tau_{\text{lep}}\tau_{\text{had}}$ channel. The green and yellow bands stand for the $\pm 1, 2\text{-}\sigma$ variations on the expected limits, respectively.

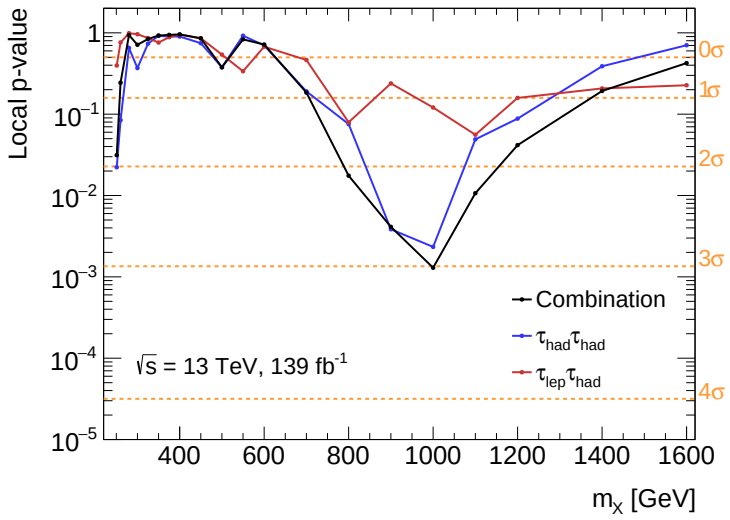


Figure 4.44: P-value of the background-only hypothesis in the search for resonant Higgs boson pair production via a narrow-width scalar particle (X) as a function of m_X. The markers indicate that the p-values derived based on the signal templates generated by simulated samples with the corresponding mass. The intermediate p-values are linearly interpolated. The black, blue and red lines corresponds to the results in the combination, the standalone $\tau_{\text{had}}\tau_{\text{had}}$ channel and the standalone $\tau_{\text{lep}}\tau_{\text{had}}$ channel. The dashed orange lines stand for the converted significance of local excess up to 4σ.

4.8 Constraints on Higgs boson self-coupling strength

The search for non-resonant HH is extended to constrain the Higgs boson self-coupling strength λ. By convention, the constraints are set on the coupling modifier $\kappa_\lambda = \frac{\lambda}{\lambda_{SM}}$ instead of λ. Only the ggF production mode is considered. The same ME generator of SM ggF HH ($\kappa_\lambda = 1$) is used to generate ggF HH with different κ_λ. Three base samples, $\kappa_\lambda = \{0, 1, 20\}$, are produced at generator level. Samples with an arbitrary value of κ_λ can be represented by the linear combination of these three samples, as shown in Equation (1.30). Because m_{HH} distribution is particularly sensitive to the variation of κ_λ. Weights are derived at the generator level in the form

$$w(m_{HH}) = \left.\frac{\Delta N(\kappa_\lambda = x)}{\Delta N(\kappa_\lambda = 1)}\right|_{m_{HH}}, \tag{4.39}$$

which is the ratio of the number of events in the linear combined sample of $\kappa_\lambda = x$ to the events in the SM samples in a certain interval ($10\ GeV$) as a function of m_{HH}. At reconstruction level, only the simulated SM and $\kappa_\lambda = 10$ ggF HH samples are available to date. The non-SM κ_λ signal events that are used in this analysis are represented by weighting the SM events. Uncertainty due to limited size of simulated samples in the derivation of the weights and the non-closure of the weighting are considered. The latter is estimated by comparing the weighted and generated $\kappa_\lambda = 10$ events, which is 3.7% on the total signal yield.

The non-SM signal events must pass the same selection requirements as the SM events. Fig-

ure 4.45 depicts the $\mathscr{A} \times \epsilon$ for signals with various values of κ_λ. It reaches a maximum of 4.7% around $\kappa_\lambda = 2$, while the lowest value is 1.6% around $\kappa_\lambda = 5$. The non-resonant HH BDT and the optimised binning are reused to extract the ggF HH signals with different κ_λ's. It is found that it can efficiently select these signals. No dedicated optimisation is considered so far. In Figure 4.46, the signals are overlaid in the same post-fit background (see Figure 4.27) BDT score distribution. Upper limits are calculated based on the same method for various κ_λ hypotheses, which is shown in Figure 4.47. The theoretical prediction of the ggF HH cross section and its uncertainty as functions of κ_λ are discussed in Section 1.2.2, which is shown in red line and error band. The κ_λ hypotheses where the theoretical predicted cross section is larger than the observed upper limit are excluded at 95% CL. In other words, the coupling modifier κ_λ, which stands for the strength of Higgs boson self coupling, is constrained at 95% CL to the interval $[-2.5, 9.7]$. The expected interval is $[-2.3, 9.8]$.

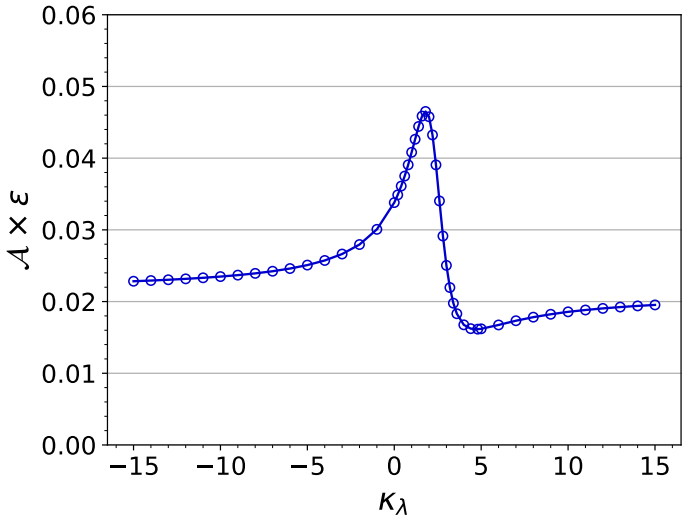


Figure 4.45: Analysis acceptance times selection efficiency of the same signal region selections as SM HH case, simulated with weighted samples with various values of κ_λ.

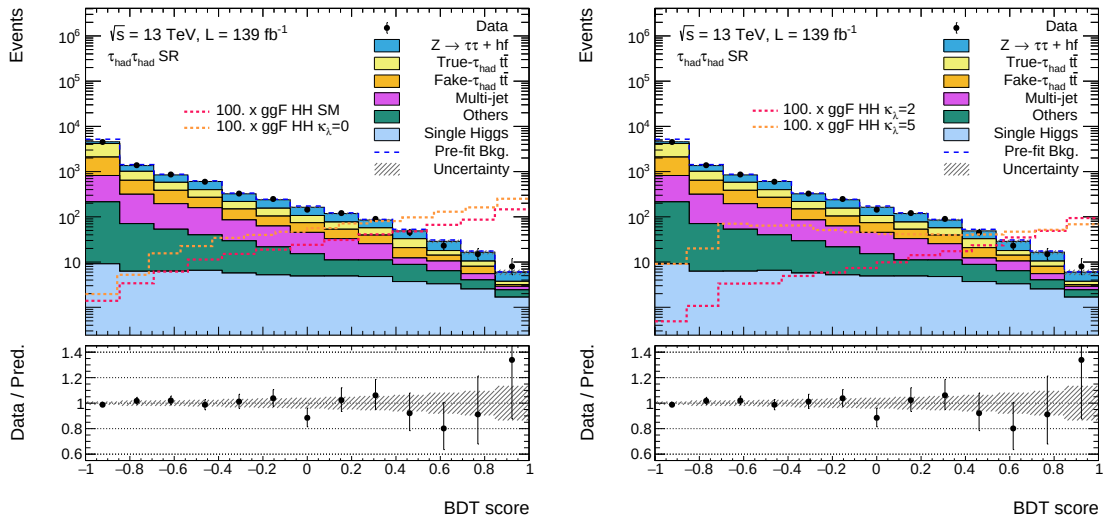


Figure 4.46: Distributions of the non-resonant HH BDT score evaluated with the SM and non-SM ($\kappa_\lambda = \{0, 2, 5\}$) ggF HH events after the background-only-hypothesis likelihood fit. The signals are scaled by a factor of 100 with respected to their cross section. In the lower panel, the ratio of data and the total background prediction is displayed in black dots. The hatched area indicates the total systematic uncertainty of the backgrounds.

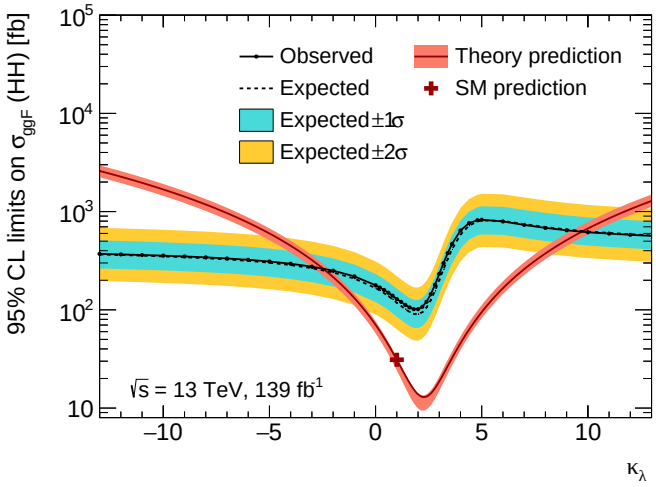


Figure 4.47: Observed (solid line) and expected (dashed line) 95% CL upper limits on the cross section of $gg \to HH$ for various κ_λ hypotheses. The markers indicate that the upper limits are calculated based on the signal templates constructed from the simulated SM HH samples or weighted from that. The intermediate upper limits are linearly interpolated. The green and yellow bands stand for the $\pm 1, 2$-σ variations on the expected limits, respectively.

Chapter 5

Search for Higgs Boson Pair Production in Boosted Topology

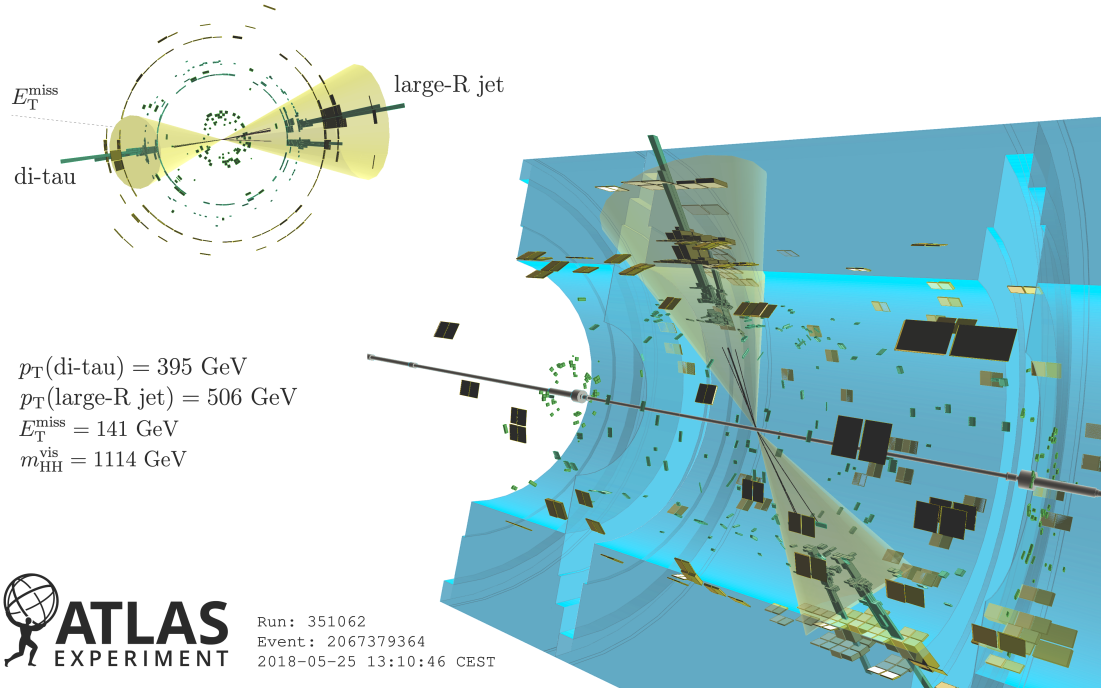


Figure 5.1: Event display of a selected boosted $X \to HH \to b\bar{b}\tau^{+}_{\mathrm{had}}\tau^{-}_{\mathrm{had}}$ candidate event.

Boosted $H \to bb$ and $H \to \tau\tau$ objects are used to extend the mass search range of resonant $X \to HH$ production up to 3 TeV. In this scenario, both Higgs bosons produced by the heavy X particle are highly Lorentz boosted. Consequently, the decay products of each Higgs boson cannot be resolved as two b-jets or two τ_{had}'s by the detector. Instead, they need to be reconstructed as a doubly b-tagged large-radius jet and a boosted di-τ object, as introduced in Section 2.3.3 and Section 3.4. Only the $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$ channel is considered in this study. It is the first analysis that uses boosted di-τ objects in ATLAS experiment. In terms of timing, this work happens after Ref. [9]

(predecessor of the resolved analysis in this thesis) and before the studies in Chapter 4. However, it is written after the resolved analysis for logical flow. This work is published in [130].

5.1 Data and simulated samples

This analysis is also based the 139 fb^{-1} LHC Run-2 p-p collision datasets. The simulated samples are generated in the same way as the resolved analysis, and the majority of the simulated samples in use are identical to those in the resolved analysis. Here only the differences are listed, the rest can be found in Section 4.1 and a general description of simulated samples in ATLAS experiments in Section 2.4.

This analysis is designed only for searching the resonant $X \to HH$ production, where the Higgs pairs are produced by narrow-width scalar resonance particles that are produced by ggF mode. The generated mass hypotheses of the BSM resonant signal include

$$1.0, 1.1, 1.2, 1.4, 1.6, 1.8, 2.0, 2.5, 3.0.$$

The unit is TeV. As mentioned in Section 4.1, these samples are simulated with FullSim.

The same SM background processes are considered as in the resolved analysis. They are W/Z + jets, multi-jet, top-quark, diboson and single Higgs boson. Multi-jet is estimated by a fake factor method which is detailed in Section 5.3.2, the rest is simulated by the same simulated samples as resolved analysis except for diboson. Top-quark production processes are found to be negligible in the signal region, whose treatment will be discussed in Section 5.3.4. Different samples are used in the boosted analysis for diboson processes compared with the resolved analysis in order to have more statistics. In preference to SHERPA 2.2, the diboson events are generated using POWHEG BOX V2 using CT10 PDF set and interfaced with PYTHIA 8 for parton shower and hadronisation.

5.2 Signal region selections

The size of data is much smaller in the boosted analysis as we are focusing on Higgs bosons with high p_{T}. Hence, unlike the resolved analysis, cut-based approach is used to extract the signal candidates in the boosted analysis, and different triggers are used as well.

5.2.1 Trigger selections

A signal event must contain two large-R jets, one is the seed jet that is further reconstructed and identified as a boosted di-τ object [1], the other one is the large-R jet that is doubly b-tagged which

[1] In this chapter, di-τ always means a fully-hadronically decaying boosted di-τ object.

arises from the $H \to bb$ decay. Thus, large-R jet triggers with lowest HLT trigger thresholds are utilised [2], which have a transverse energy (E_{T}) thresholds from 360 to 460 GeV, depending on the data-taking period.

Similar as in the resolved analysis, an offline requirement on the large-R jet p_{T} with a value of 40-50 GeV larger than the online trigger threshold is applied in order to ensure a constant trigger efficiency. The offline p_{T} selection is applied on the leading p_{T} large-R jet, regardless of whether this jet is matched to the boosted $b\bar{b}$ or $\tau_{\mathrm{had}}\tau_{\mathrm{had}}$. The online/offline selection criteria and the corresponding data=taking periods are summarised in Table 5.1.

Table 5.1: Online HLT large-R jet trigger E_{T} threshold and the corresponding offline p_{T} selection, according to the data-taking period.

Period	Online E_{T} threshold [GeV]	Offline p_{T} selection [GeV]
2015, 2016	360	>410
2016	400	>450
2016, 2017	420	>460
2016, 2017	440	>480
2017, 2018	460	>500

5.2.2 Event selections

The events that pass the trigger selections are further selected. The observables are defined by the objects described in Section 2.3 and for boosted di-τ in Section 3.4. Firstly, the events are vetoed if they contain an electron or muon that reach the quality defined in Table 2.3. Also, they are required to contain at least one reconstructed boosted di-τ object. In addition to the basic selections on di-τ described in Section 3.4, the di-τ candidates also need to fulfil the following requirements:

- The number of sub-jets of di-τ is less or equal to three;
- The two p_{T}-leading sub-jets are required to have $p_{\mathrm{T}} > 50$ GeV to reject the contamination of multi-jet background;
- The angular distance ΔR between the two p_{T}-leading sub-jet is less than 0.8 to insure they are fully contained in the same seed jet;
- The electric charges of the two p_{T}-leading sub-jets have opposite signs (OS);
- The p_{T} of di-τ is larger than 300 GeV [3], which is calculated from the two p_{T}-leading sub-jets, and the pseudorapidity satisfies $|\eta| < 2.0$ (not 1.37–1.52).

Besides, a large-R jet which corresponds to the $H \to bb$ decay is required to have $p_{\mathrm{T}} > 300\ GeV$ [4], $|\eta| < 2.0$ and a combined mass $m > 50$ GeV. The large-R jet must separate from the selected di-τ by

[2] There are no dedicated boosted di-τ triggers implemented in Run2.

[3] Or the trigger related offline p_{T} selection if the jet that fires the trigger is matched to the seed jet of the di-τ object.

[4] Or the trigger related offline p_{T} selection if the jet that fires the trigger is matched to the seed jet of the di-τ object.

a distance of $\Delta R > 1.0$ to avoid double counting of objects. If there are multiple di-τ or large-R jet candidates, for both objects the one with largest p_{T} is selected.

Events are required to have a $E_{\mathrm{T}}^{\mathrm{miss}}$ greater than 10 GeV, which is used to define the direction of the missing transverse momentum and to reject the contribution from the multi-jet background. The selected di-τ object needs to pass the 60% efficiency identification working point, and the selected large-R jet must be doubly b-tagged (see Table 2.4). Events are removed if there are more than one 2-b-tagged large-R jets. This criterion is applied to insure orthogonality to the searches in $b\bar{b}b\bar{b}$ final state.

To further extract the signal, the azimuthal angle between the selected di-τ and the missing transverse momentum ($|\Delta\phi(\text{di-}\tau, E_{\mathrm{T}}^{\mathrm{miss}})|$) is required to be less than 1.0 in order to further reduce the contamination of multi-jet background and to define a control region enriched in multi-jet events. And the combined mass of the large-R jet must lie between 60 and 160 GeV. Finally, different thresholds are applied to the visible invariant mass of the HH system to improve the sensitivity for higher mass hypotheses, which are 0.9 TeV for mass points from 1.6 to 2.0 TeV and 1.2 TeV for mass points from 2.5 to 3.0 TeV. The above object-level and event-level selections are summarised in Table 5.2. The cumulative analysis acceptance and selection efficiency $\mathscr{A} \times \epsilon$ at each stage of the selection chain is estimated with simulated resonant signal samples with various mass hypotheses in Table 5.3. The limiting factors are the high online and offline p_{T} threshold of di-τ and large-R jet for the lower mass signals as the Higgs bosons are less boosted, and the reconstruction and identification efficiency of di-τ tagging and double b-tagging for the higher mass signals as the underlying tracks become difficult to resolve. The final expected $\mathscr{A} \times \epsilon$ varies from 0.29% to 3.2% depending on the mass, as shown in Figure 5.2. In the end, the expected number of events is much lower than the resolved analysis. Prior to the m_{HH} selection, only 1.4 background event is expected (see Table 5.8 in the next subsection). Therefore, the number of total events is used for the final signal extraction, i.e., counting experiment with a single bin.

As what is done in the resolved analysis, CRs and VRs are defined by varying the selection criteria of the SR, in order to estimate or validate a specific process. The definition of these regions are followed in Section 5.3 where the predictions on background processes are discussed.

5.3 Background estimation

With the selection criteria, especially the requirement of one di-τ and one large-R jet with high transverse momentum, the background contamination becomes much lower compared with resolved analysis. It is found that $Z \to \tau\tau + hf$ and multi-jet are the main background processes in the signal region. The methodology of estimating them are detailed in the following subsections (Section 5.3.1 and 5.3.2) respectively, while the other background processes are modelled by MC simulation. Besides, the scale factor of true boosted di-τ tagging efficiency is evaluated as part of this analysis, which is described in Section 5.3.3.

Table 5.2: Summary of object-level and event-level selections that defines the SR of $HH \to b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ boosted analysis. $N(X)$ denotes the number of object X.

Variable	Requirement
Di-τ	
p_{T}	> 300 GeV (or 410-500 depending on trigger)
$\lvert\eta\rvert$	$\lvert\eta\rvert < 2.0$ (not 1.37–1.52)
N(sub-jet)	≤ 3
Two p_{T}-leading sub-jets)	> 50 GeV
	opposite-sign electric charge
	$\Delta R < 0.8$
large-R jet	
p_{T}	300 (or 410-500 depending on trigger)
$\lvert\eta\rvert$	$\lvert\eta\rvert < 2.0$
Combined mass	> 50 GeV
ΔR(di-τ, large-R jet)	> 1.0
Event-level	
Trigger	Pass and di-τ or large-R jet trigger-matched
$N(e/\mu)$	$= 0$
N(di-τ)	$= 1$, pass 60% identification working point
N(large-R jet)	$= 1$, two b-tagged
$E_{\text{T}}^{\text{miss}}$	> 10 GeV
$\lvert\Delta\phi$(di-τ, $E_{\text{T}}^{\text{miss}})\rvert$	< 1.0
$m_{\text{large-}R\text{ jet}}$	$60 - 160$ GeV
m_{HH}	>0, 0.9, 1.2 TeV depending the signal mass hypothesis

Table 5.3: Cumulative analysis acceptance times selection efficiency at each stage of the selection chain estimated using the simulated resonant ($X(m_X$ in TeV)) signal samples. The efficiencies are calculated with respect to $HH \to b\bar{b}\tau^+_{\text{had}}\tau^-_{\text{had}}$ events. Due to technical reason the selections are not applied in the natural way. 'Di-τ selection' includes all requirements on di-τ and its sub-jets except the opposite-sign requirement on the charges of the leading two sub-jets, which is included in 'other event-level selections'

Selection stage	$X(1.0)$	$X(1.2)$	$X(1.6)$	$X(2.0)$	$X(3.0)$
Trigger and object definitions at pre-selection level	24%	40%	56%	61%	53%
Di-τ selections	5.0%	9.9%	16%	17%	9.6%
Trigger-dependent p_{T} requirements	1.7%	7.3%	14%	17%	9.5%
Large-R jet selections	1.4%	6.2%	13%	15%	8.6%
Doubly b-tagging of the large-R jet	0.43%	2.1%	3.9%	4.1%	1.9%
Other event-level selections	0.28	1.6%	2.9%	3.1%	1.5%

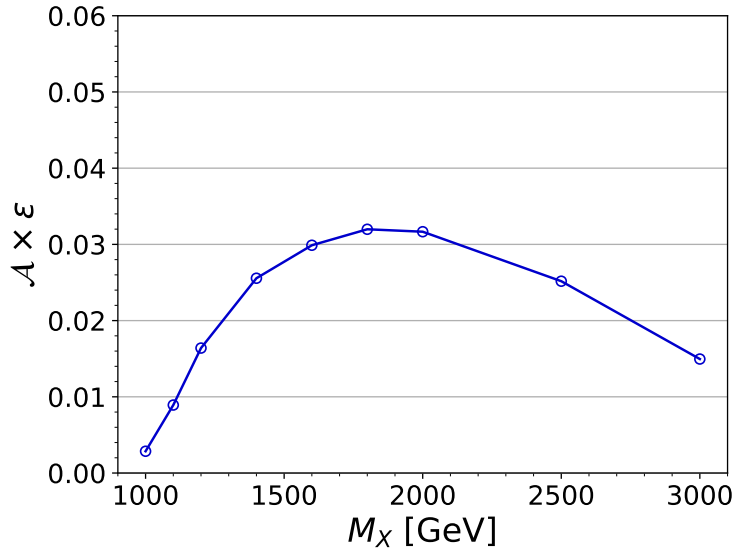


Figure 5.2: Analysis acceptance times selection efficiency of the full signal region selections estimated with simulated signal samples as a function of the mass of the resonant particle m_X.

5.3.1 Z+heavy flavour jets

Similar as the resolved analysis, $Z \to \tau\tau + hf$ is also the largest and irreducible background in the boosted analysis. Using the same generator level parton matching procedure is employed, splitting $Z \to \tau\tau$ + jets into $Z \to \tau\tau + bb, bc, cc, bl, cl, ll$. Here we define heavy flavour as $hf = bb, bc, cc, bl$ and light flavour as $lf = cl, ll$. As discussed, there is a known discrepancy between the number of $Z + hf$ events in the prediction of SHERPA 2.2 simulated sample and in data. Instead of including the $Z + hf$ CR in the likelihood function, the data-driven normalisation correction factor μ_{Z+hf} is measured in a $Z + hf$ CR separately, which is,

$$\mu_{Z+hf} = \frac{N(\text{data}) - N(\text{MC, not } Z + hf)}{N(\text{MC}, Z + hf)}, \tag{5.1}$$

where N is the number of events. Non-$Z + hf$ are subtracted from data and the subtracted event yield is compared with the prediction by simulated $Z + hf$ samples.

Following the same idea as the resolved analysis, the $Z + hf$ CR is defined in $b\bar{b}\ell^+\ell^-$ ($\ell = e/\mu$) final state. Events in this region are triggered by one of the single electron triggers and single muon triggers listed in Table 4.6, where the online and offline p_{T} thresholds are shown. The di-lepton system is required to have an invariant mass $81\text{GeV} < m_{\ell\ell} < 101\text{GeV}$ and a transverse momentum $p_{\mathrm{T}}^{\ell\ell} > 300\text{GeV}$. These selections are analogous to the selections on the boosted di-τ object in the SR. Also, the events must have at least one 2-b-tagged large-R jet with mass $m_J > 50\text{GeV}$ and transverse momentum $p_{\mathrm{T}}^{J} > 250\text{GeV}$. After the event selections, the fraction of $Z \to \ell\ell + hf$ events is about 80%. Table 5.4 summarises the event yields in the $Z + hf$ CR.

Table 5.4: Event yields of data and background prediction in the $Z + hf$ CR.

Process	Yield ±MC Stat. Unc.
$Z \to \ell\ell + hf$	150 ±2.5
$Z \to \ell\ell + lf$	6.8 ±0.95
Diboson	27 ±0.76
$VH(H \to bb)$	11 ±0.04
Others (W+jets, top-quark)	1.1 ±0.47
Total Background	200 ±2.8
Data	230

The normalisation factor is calculated based on the yields. It is

$$\mu_{Z+hf} = 1.2 \pm 0.10(\text{stat.}) \pm 0.28(\text{syst.}), \tag{5.2}$$

where the statistical uncertainty (stat.) accounts the limited size of the data and simulated samples, dominantly from data. The systematic uncertainty (syst.) arises from the experimental uncertainties [5] and the extrapolation uncertainty from the $Z + hf$ CR to the SR, dominantly from the latter, which is estimated by the difference between the p_{T} distribution of the visible decay products of $Z \to \ell\ell$ and $Z \to \tau\tau$ events. Comparison between data and the $Z + hf$ prediction after applying the normalisation factor is shown in Figure 5.3, with the $p_{\mathrm{T}}^{\ell\ell}$ distribution. μ_{Z+hf} is found to be compatible with the normalisation factor of $Z + hf$ in the resolved analysis (see Section 4.3.1) and in a $VH(H \to bb)$ analysis using boosted $H \to bb$ topology [175] (1.32 ± 0.16), therefore it can be applied in the SR with confidence despite the lack of dedicated $Z \to \tau\tau + hf$ validation regions.

5.3.2 Multi-jet

As shown in Section 3.4, The di-τ identification algorithm is very efficient at rejecting quark or gluon initiated jets against multi-jet events. However, there is an approximately probability of 10^{-4} for a jet to be misidentified as a di-τ. The misidentified jet is called a fake di-τ in the following descriptions. As a consequence of the large cross section of the multi-jet process, the contamination of multi-jet events that contain fake di-τ is not negligible.

Similar as the fake-τ_{had} multi-jet in the resolved analysis, the contribution of fake-di-τ multi-jet in boosted analysis is also estimated by the fake factor (FF) method. A region with a zero b-tagged large-R jet and same-sign (SS) of the electric charge of the two leading sub-jets in the di-τ are selected, which is called a fake factor control region (FF-CR). It is enriched in multi-jet events and it has enough data statistics to derive FF. The other selections are same as those for the signal region, but the mass window selection (60-160 GeV) on the mass of the large-R jet (m_J) and the selections on m_{HH} are not applied in the FF-CR. Besides, analogous to the FF method in the

[5] Same as the uncertainties considered in the signal region (see Section 5.4) and the uncertainty for electron, muon (see Section 4.4.1) and small-R jet (reduced version of the terms listed in Table 4.10.)

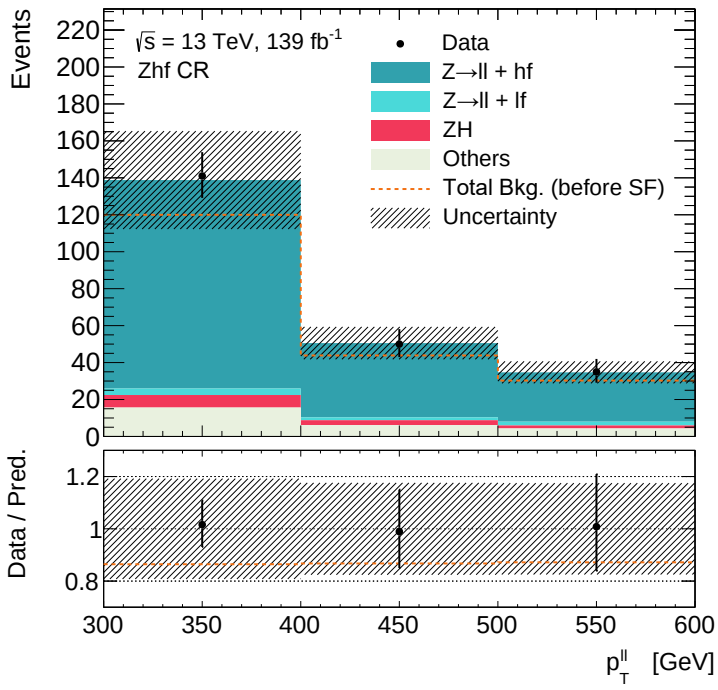


Figure 5.3: Comparison of the distribution of $p_{\mathrm{T}}^{\ell\ell}$ between data and background prediction after applying the normalisation factor of $Z + hf$ in the $Z + hf$ CR. The ratio is shown in black dots in the lower panel. The total background before applying the normalisation factor is displayed in the dashed red line. The uncertainty shown in hatched bands includes all the experimental uncertainties and the uncertainties in the normalisation factor.

resolved analysis, the ID region is defined by requiring the di-τ to pass the identification [6], while the Anti-ID region is defined by requiring the di-τ to fail the identification but pass a very loose working point [7]. The di-τ in the anti-ID region is call an anti-di-τ. Given the region defined, the fake factor is derived by Equation (4.7) as a function of di-τ p_{T}, as shown in Figure 5.4. It is found that there is no significantly dependency for the FF on the other variables, e.g., the pseudorapidity of di-τ. The subtracted non-multi-jet processes contributes about 24% and 3% of the total event yields in the numerator and the denominator, respectively.

Systematic uncertainties are taken into account for different selections in the FF-CR where the FF is calculated, and the SR (or VRs) where the FF is applied. One is introduced to account the extrapolation from same-sign charges of di-τ sub-jets selection to opposite-sign selection, and another one to account the extrapolation from 0-b-tagged large-R jet selection to 1- and 2-b-tagged selections. They are derived from the difference between the FF calculated in the regions with alternative selections. Another source of uncertainty is due to the statistical uncertainty of data and the uncertainty of the subtracted simulated non-multi-jet samples, dominantly from the limited statistics of the data. The overall impact of the FF uncertainties on the multi-jet background is about 12%–62%, depending on the p_{T} of di-τ.

[6] 60% efficiency working point.
[7] 99% efficiency working point.

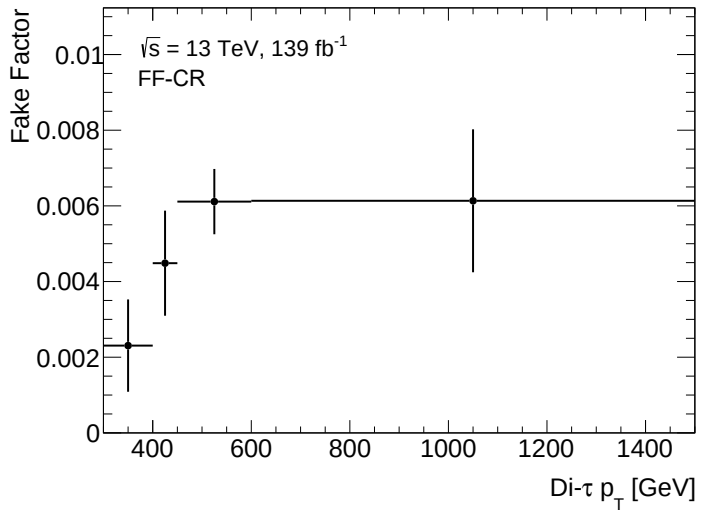


Figure 5.4: Fake factors as a function of the p_{T} of the di-τ objects. The error bar indicates the uncertainty of the FF that originates from the statistical uncertainty of data and the uncertainty of the subtracted simulated non-multi-jet samples, dominantly from the limited statistics of the data.

The FF is validated in zero and one b-tagged regions (FF-OS-VR) whose definitions are outlined in Table 5.5, where the different selections with respect to the FF-CR are shown. For the VRs and SR, the number of multi-jet events ($N_{\text{ID region}}$) is predicted by

$$N_{\text{ID region}} = FF \times \left[N_{\text{Anti-ID region}}(\text{data}) - N_{\text{Anti-ID region}}(\text{non-multi-jet})\right], \tag{5.3}$$

where 'Anti-ID region' stands for the corresponding anti-ID region. Comparisons between data and background predictions on the distributions of m_{HH} are shown in Figure 5.5. In 1-b-tagged FF-OS-VR, we only count the total number of events due to the limitation of statistics. Good agreement is observed in 0-b-tagged FF-OS-VR. Considering the discrepancy observed in 1-b-tagged FF-OS-VR, an additional non-closure uncertainty of 50% is assigned for 1- or 2-b-tagged multi-jet events.

Table 5.5: Definitions of the validation regions of the fake factor in the boosted analysis. The last two selections in 1-b-tagged FF-OS-VR are intended to eliminate the contamination of signal events.

Region name	Selections different from FF-CR
0-b-tagged FF-OS-VR	Sub-jets of di-τ have opposite-sign charges $\lvert\Delta\phi(\text{di-}\tau, E_{\mathrm{T}}^{\mathrm{miss}})\rvert > 1.0$
1-b-tagged FF-OS-VR	Sub-jets of di-τ have opposite-sign charges large-R jet is 1-b-tagged $\lvert\Delta\phi(\text{di-}\tau, E_{\mathrm{T}}^{\mathrm{miss}})\rvert > 1.0$ Mass of large-R jet $m \in (50, 60) \bigcap (160, \infty)$ GeV (mass sideband) $m_{HH} < 1500$ GeV

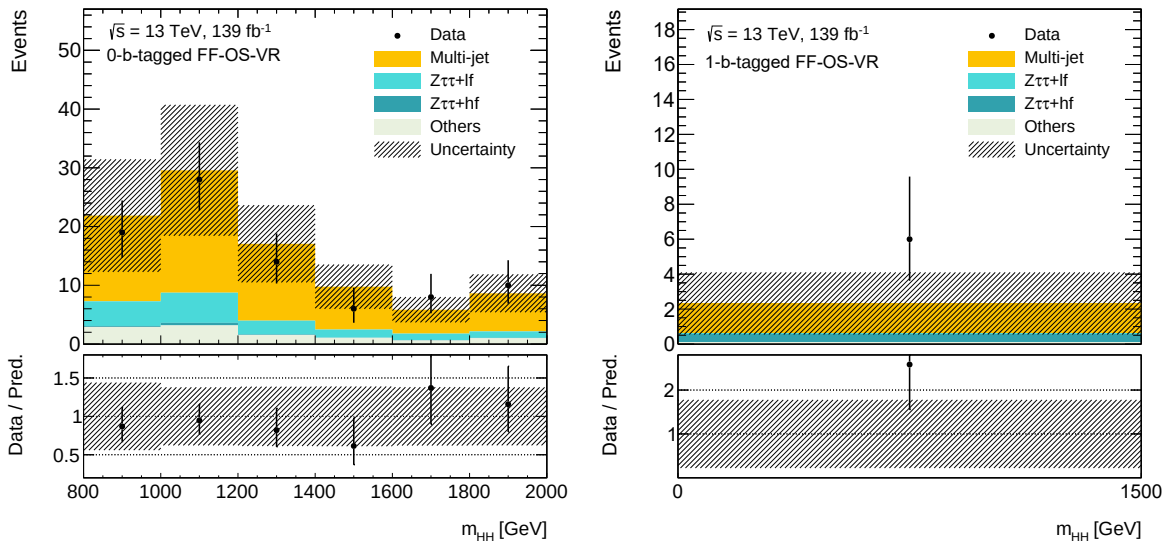


Figure 5.5: Distribution of the visible invariant mass of the HH system in the 0- and 1-b-tagged FF-OS-VR. The x-axis of the figure for 1-b-tagged FF-OS-VR indicates the selection applied on m_{HH}. The data is compared with the background prediction, where the multi-jet is predicted by the FF method. The ratio is shown in black dots in the lower panel. The uncertainty shown in hatched bands includes all the experimental uncertainties and the FF uncertainties.

5.3.3 Calibration of di-τ tagging

For the di-τ tagger to be applicable to the physics analysis, the tagging efficiency on the di-τ that is matched to generator level true di-τ (truth-matched) in simulated samples needs to be corrected by measuring the response on the di-τ in real data. A dedicated region enriched with $Z \to \tau_{\text{had}}\tau_{\text{had}}$ events that contain boosted di-τ objects, labelled as $Z\tau\tau$ CR, is defined to measure the scale factor that quantify the difference on tagging efficiency between di-τ in read data and in simulation. It is

$$SF = \frac{N(\text{data}) - N(\text{di-}\tau \text{ is not truth-matched})}{N(\text{di-}\tau \text{ is truth-matched})}, \tag{5.4}$$

where N stand for the number of events [8]; N(di-τ is not truth-matched stands for the events in which the di-τ is from mis-identification, they are subtracted from the data. The scale factor is applied to the truth-match di-τ in all background and signal processes.

The $Z\tau\tau$ CR is defined based on the SR selections, but without the mass window selection on m_J and the mass thresholds on m_{HH}. On top of that, it selects a 0-b-tagged large-R jet and require the visible di-τ mass [9] to lie between 30 and 90 GeV to get a higher purity of truth-matched di-τ events in this region.

No significant dependences of the SF and its uncertainties on the kinematics of di-τ are found,

[8] Note that there is one di-τ per event.

[9] The visible mass of di-τ equals to the invariant mass of the two p_{T}-leading sub-jets in the di-τ.

therefore the *SF* is a single number with overall uncertainties, which is

$$SF = 0.84 \pm 0.09(\text{stat})^{+0.14}_{-0.13}(Z\text{+jets modelling})^{+0.19}_{-0.20}(\text{syst}), \tag{5.5}$$

where the following sources of uncertainties in the *SF* are considered:

- Statistical uncertainty due to the limited size of the data and simulated samples (stat);
- The theoretical uncertainty of the Z+jets processes (Z+jets modelling). The uncertainty on the cross section is 5%. The uncertainty on $\mathscr{A}\times\epsilon$ is derived in the same approach as described in Section 4.4.2.2 for the $Z + hf$ in the resolved analysis, which is found to only impact the normalisation of Z+jets.
- The other uncertainties, including the experimental uncertainties (see Section 5.4) and FF uncertainties in multi-jet process (syst). The theoretical uncertainties in other processes are found to be negligible.

It is only different from the FF-CR in the selection of $|\Delta\phi(\text{di-}\tau, E_{\text{T}}^{\text{miss}})|$.

Figure 5.6 shows the data and background prediction comparisons in the distribution of various di-τ kinematic observables in the $Z\tau\tau$ CR. The total background before applying the *SF* is displayed in the dashed red line. Good agreement is achieved after the *SF* is applied. The *SF* is validated in a 1-*b*-tagged validation region (1-*b*-tagged $Z\tau\tau$ VR), this region is same as the 1-*b*-tagged FF-OS-VR but with an opposite selection of $|\Delta\phi(\text{di-}\tau, E_{\text{T}}^{\text{miss}})|$, to select a $Z \to \tau_{\text{had}}\tau_{\text{had}}$ enriched region and to reject multi-jet events. Due to the limitation of the sample size, only the event number counting is compared, as shown in Figure 5.7. Good agreement is observed in this region, giving the confidence to apply it in 2-*b*-tagged regions.

5.3.4 Top-quark background

The $t\bar{t}$ process in the boosted analysis is not as contributory as in the resolved analysis. This can be explained by two facts. Firstly, the two τ_{had} candidates and two b-jet candidates are less unlikely to be close to each other in a $t\bar{t}$ event. Which is predicted in Figure 4.3, where the $\Delta R(\tau_{\text{had}}, \tau_{\text{had}})$ and $\Delta R(b, b)$ distributions of $t\bar{t}$ events (majority of the background labelled as 'Top-quark') result in a most probable value of about 3. Secondly, the di-τ identification algorithm is very efficient in rejection of $t\bar{t}$ background, as illustrate in Figure 5.8. The events shown are before applying the selection on the 60% efficiency di-τ identification working point (corresponding to a di-τ BDT score of 0.72), as well as no requirement on the number of *b*-tagging of large-*R* jet, $\Delta\phi(\text{di-}\tau, E_{\text{T}}^{\text{miss}})$, m_J mass window or m_{HH}. A good rejection of $t\bar{t}$ background is predicted and the single top-quark background is found to be negligible.

Using the largest available $t\bar{t}$ samples, there are still no raw MC events passing the SR selections. In order to account for possible contribution from $t\bar{t}$, a $+1\sigma$ uncertainty of 0.12 events in introduced to the event yield of $t\bar{t}$ process. The number is estimated by the number of remaining

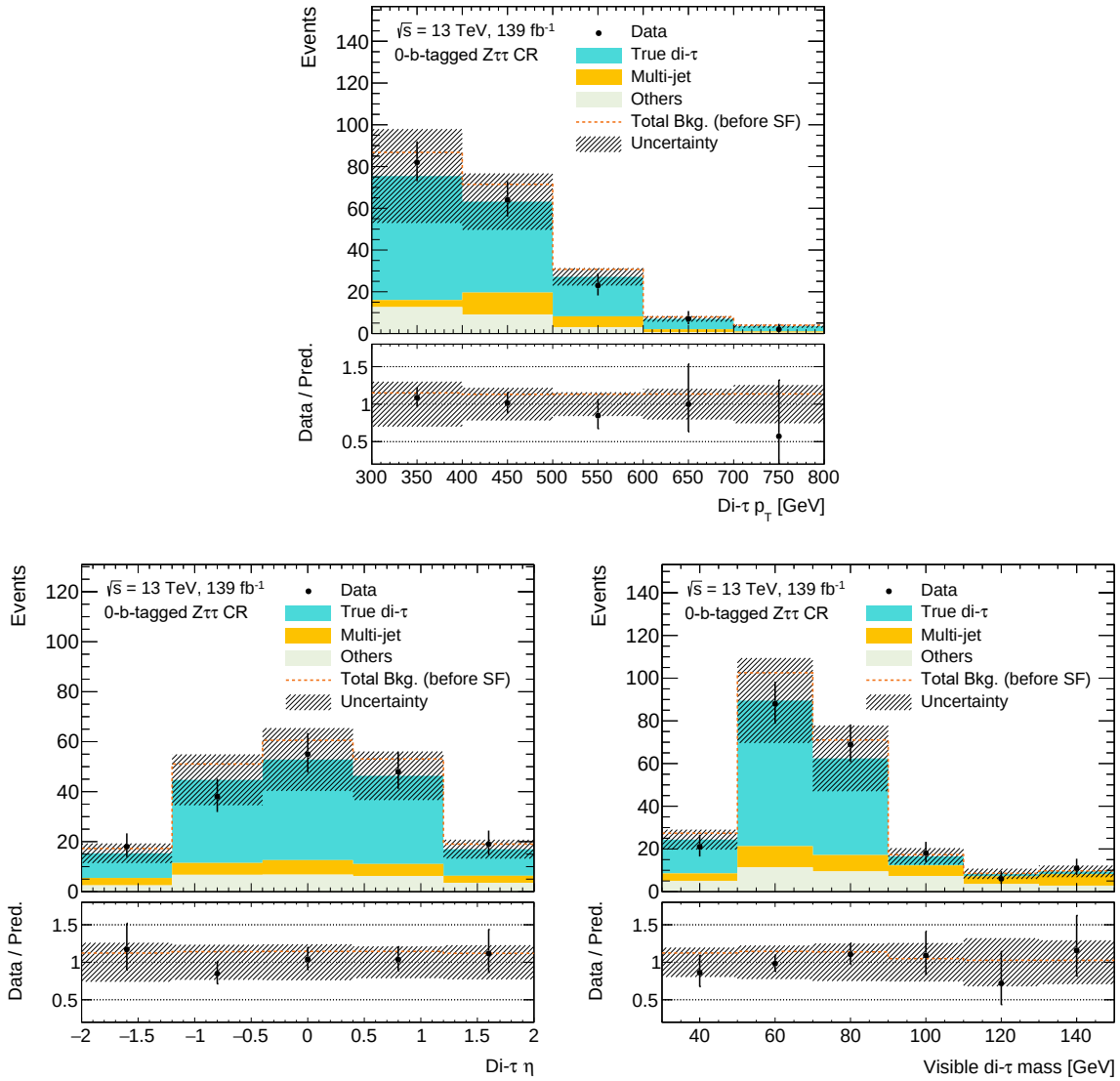


Figure 5.6: Distribution of the p_{T}, η and visible mass of di-τ after applying the *SF* in the $Z\tau\tau$ CR. Data is compared with the background prediction in the ratio shown in black dots in the lower panel. The total background before applying the *SF* is displayed in the dashed red line. The uncertainty shown in hatched bands includes all the experimental uncertainties and the uncertainties in *SF*.

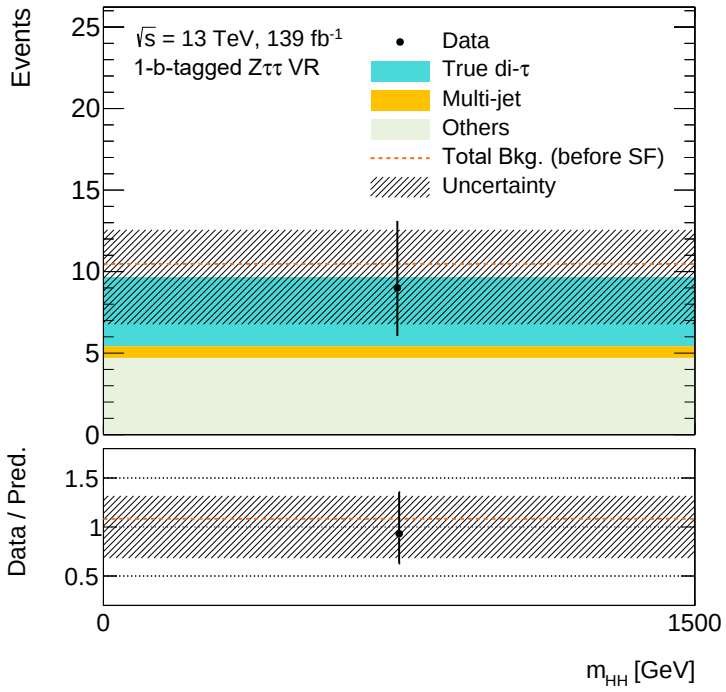


Figure 5.7: Comparison of the counting of event number between data and background prediction after applying the *SF* in the 1-*b*-tagged $Z\tau\tau$ VR. The main contribution to the background labelled as 'Others' is the $t\bar{t}$ process (about 3.6 events). The ratio is shown in black dots in the lower panel. The total background before applying the *SF* is displayed in the dashed red line. The x-axis indicates the selection applied on m_{HH}. The uncertainty shown in hatched bands includes all the experimental uncertainties and the uncertainties in *SF*.

$t\bar{t}$ events prior to the m_J mass window and m_{HH} selection.

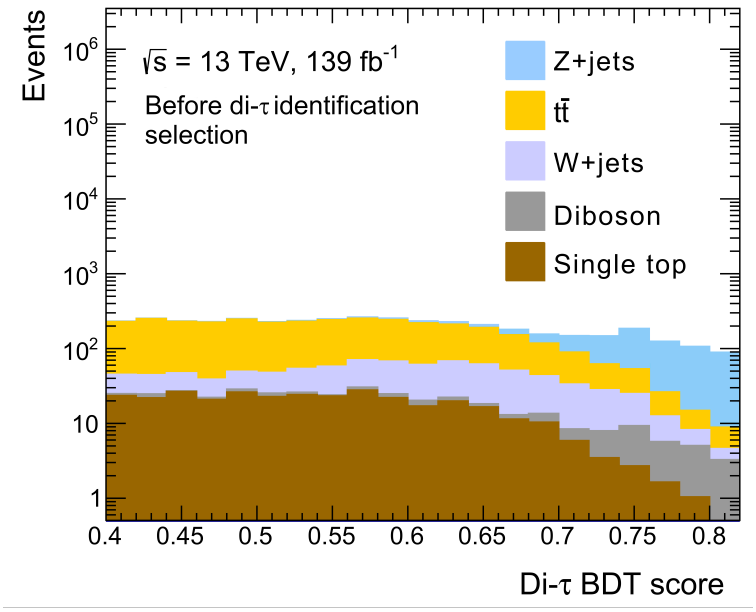


Figure 5.8: Distribution of di-τ identification BDT score on various background processes in a pre-SR selection stage. The 60% efficiency di-τ identification working point selection correspond to a score of 0.72.

5.4 Systematic uncertainties

The statistical uncertainty of the collision data is the dominating uncertainty because of the small number of observed events. The systematic uncertainties are considered and they only affect the total number of events. Similar as the resolved analysis, the statistical uncertainty of the MC simulation samples is accounted. The two leading background processes, $Z + hf$ and multi-jet, are estimated using the data-driven approach, so the uncertainties on them originate from those on the normalisation factor of $Z + hf$ and the fake factor of multi-jet, detailed in Section 5.3 and summarised in Table 5.6. The uncertainty in $\mathscr{A} \times \epsilon$ for the simulated $X \to HH$ signal samples are evaluated by the same approach as for the $X \to HH$ signal in the resolved analysis. It is 2%–4% for most of the mass hypotheses. For the other simulated processes, the uncertainty in the cross section and $\mathscr{A} \times \epsilon$ is found to be negligible. The impact of each source of uncertainties to the final results is discussed in Section 5.6.

The experimental uncertainties include those for the resolved $\tau_{\text{had}}\tau_{\text{had}}$ channel analysis that are listed in Section 4.4.1 except those for the resolved τ_{had} and small-R jet [10]. Uncertainties are also accounted for the scale and resolution of the energy and mass for the large-R jets. Each is measured in a dedicated analysis with large-R jet events in situ using the early Run-2 dataset, as reported in Ref. [100]. Moreover, the uncertainty in the di-τ tagging efficiency is measured together with the

[10] Uncertainty of small-R jet affect the energy scale of $E_{\text{T}}^{\text{miss}}$ but the impact are found to be negligible, therefore it is not included.

this analysis in Section 5.3.3. The uncertainty in the energy scale of di-τ is estimated by varying the passive material in the GEANT4 simulation and physics list in the parton shower generation using simulated samples. Table 5.7 outlines these additional experimental uncertainties with respect to Table 4.9 and 4.10.

Table 5.6: Summary of systematic uncertainties related to fake-τ_{had} $t\bar{t}$ and multi-jet estimations.

Name of uncertainty	Description
ZhfSF_Stat	Statistical uncertainty of the $Z+hf$ normalisation factor
ZhfSF_Syst	Systematic uncertainty of the $Z+hf$ normalisation factor including experimental uncertainty and the extrapolation from $Z+hf$ CR to SR
FF_Stat	Statistical uncertainty of the *FF*s
FF_Transition_Sign	SS to OS region extrapolation uncertainty
FF_Transition_Btag	0-tag to 1/2-tag region extrapolation uncertainty
FF_1btag_Sideband	1-*b*-tagged FF-VR non-closure uncertainty

Table 5.7: Summary of experimental systematic uncertainties additional to the boosted analysis.

Name of uncertainty	Description
FATJET_Medium_JET_Comb_Baseline_Kin FATJET_Medium_JET_Comb_Modelling_Kin FATJET_Medium_JET_Comb_Tracking_Kin FATJET_Medium_JET_Comb_TotalStat_Kin	Uncertainties in mass and energy scale of large-*R* jet
FATJET_JER	Uncertainties in large-*R* jet energy resolution
FATJET_JMR	Uncertainties in large-*R* jet mass resolution
TAUS_TRUEHADDITAU_EFF_JETID_TOTAL	MC-based di-τ tagging efficiency uncertainty
TAUS_TRUEHADDITAU_SME_TES_TOTAL	MC-based di-τ energy scale uncertainty
DiTauSF_Stat	Statistical uncertainty in the di-τ tagging efficiency scale factor
DiTauSF_Syst	Systematic uncertainty in the di-τ tagging efficiency scale factor including experimental uncertainty and multi-jet uncertainties
DiTauSF_ZMODEL	Systematic uncertainty in the di-τ tagging efficiency scale factor from the theoretical uncertainty of Z+jets

5.5 Statistical analysis

5.5.1 Likelihood function

The likelihood function in the boosted analysis is constructed in the same scheme as described in Section 4.5. However, there is only one counting experiment (one bin) and there are no free parameters for the normalisation of backgrounds in the boosted analysis. Consequently, Equation (4.23) can be simplified to

$$\mathscr{L}(\mu,\{\gamma,\boldsymbol{\alpha}\}) = Poisson\left(n;\gamma\left[\mu S(\boldsymbol{\alpha}) + \sum_b B_b(\boldsymbol{\alpha})\right]\right) \times Poisson(m;\gamma\tau) \times \prod_a Gaussian(\alpha_{m,a};\alpha_a,1). \tag{5.6}$$

5.5.2 Test statistics with pseudo experiments

As discussed in Section 4.5.2, the asymptotic approximation on the p.d.f of the profile likelihood ratio test statistic can only be used in its valid region. In the resolved analysis, the requirement on a

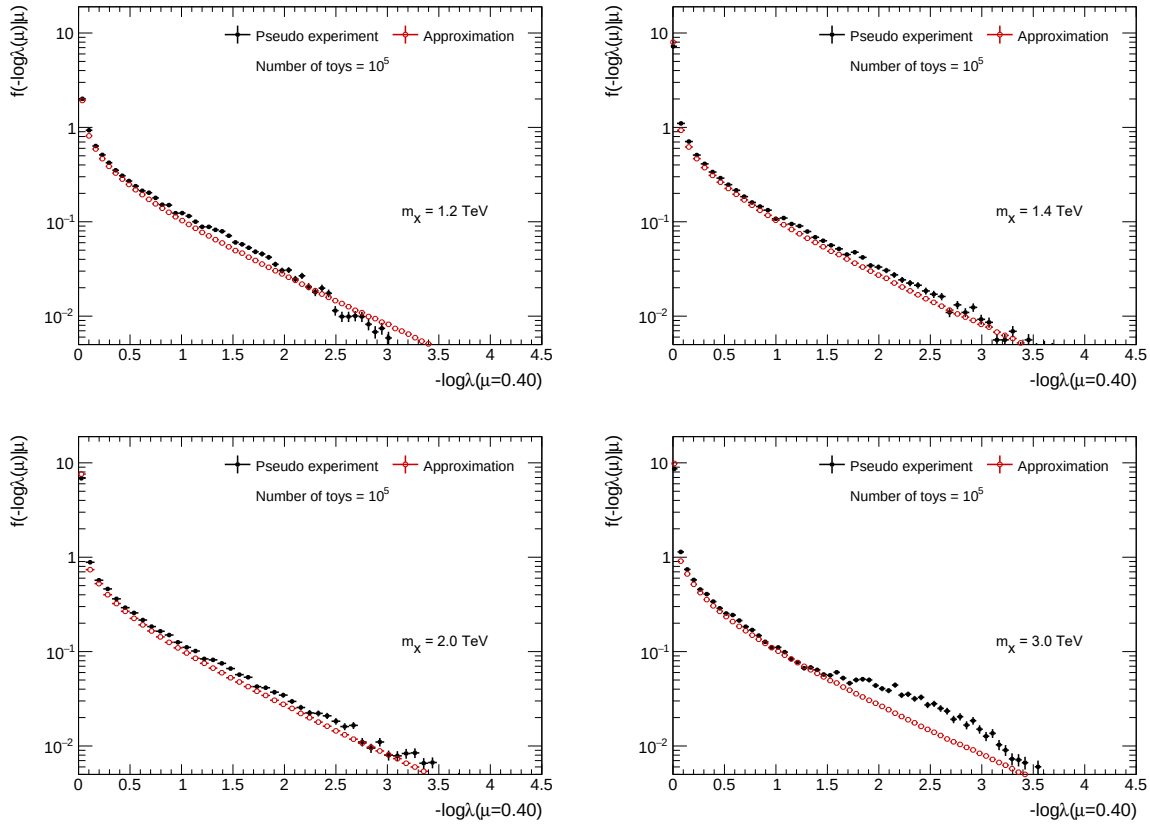


Figure 5.9: Illustration of the distribution of the profile likelihood ratio test statistic $(-log\lambda_\mu|\mu)$ sampled by pseudo-experiments in black and by asymptotic approximation (half χ^2-distribution with one degree of freedom) in red, for hypothesis tests in various mass points.

minimum expected background events of 5 in the binning optimisation algorithm insure that, while in the boosted analysis, the number of expected background events is too low. Therefore, the p.d.f of the test statistic has to be generated by pseudo experiments. A demonstration of the validation on asymptotic approximation is shown in Figure 5.9. The discrepancies support the conclusion.

5.6 Results

Table 5.8 summarise the event yields of data and expected background processes after the SR selections and the background estimations. Figure 5.10 shows the distribution of m_{HH} in the SR before the selection on that variable. Two events are observed with the value of m_{HH} at 1012 GeV and 1114 GeV. The bin edges indicate the thresholds of the m_{HH} selections.

To extract the observed and expected upper limits on the cross section of $X \to HH$, the p.d.fs of the test statistic under both signal+background and background-only hypotheses are generated at each scanning point of the POI μ, with 10k pseudo experiments. Using the sampled test statistic distributions, the same approach as introduced in Section 4.5.2 is used to extract the CL_s-based upper limit at 95% confidence-level on the cross section of resonant HH production via a narrow-width heavy scalar particle X. The results are shown in Figure 5.11, which is derived based on the

Table 5.8: Number of data and expected background events in the SR. The uncertainty includes all the theoretical, experimental and statistical uncertainties.

Selection on m_{HH}	> 0 TeV	> 0.9 TeV	> 1.2 TeV
$Z \to \tau\tau + hf$	$0.89^{+0.45}_{-0.43}$	$0.75^{+0.51}_{-0.43}$	0.17 ± 0.09
$Z \to \tau\tau + lf$	0.05 ± 0.05	0.05 ± 0.05	-
Multi-jet	0.18 ± 0.14	0.17 ± 0.13	0.09 ± 0.07
ZH	0.11 ± 0.04	0.09 ± 0.03	0.02 ± 0.01
Other backgrounds	$0.13^{+0.16}_{-0.09}$	$0.13^{+0.16}_{-0.09}$	$0.05^{+0.12}_{-0.04}$
Total background	$1.4^{+0.49}_{-0.46}$	$1.2^{+0.56}_{-0.46}$	$0.33^{+0.17}_{-0.12}$
Data	2	2	0

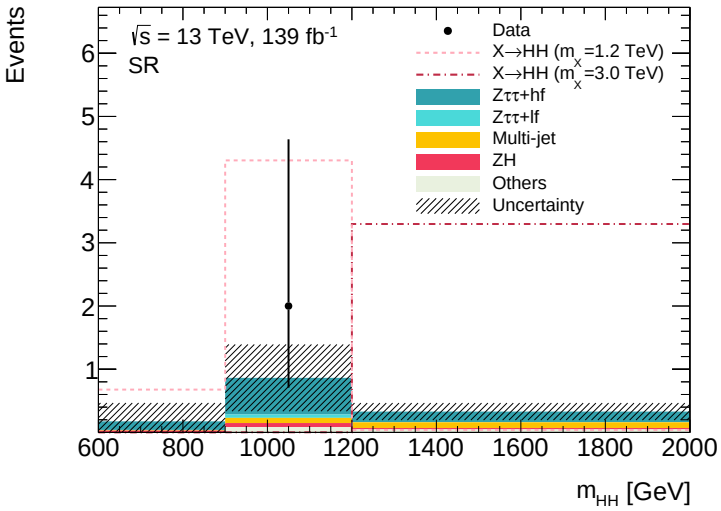


Figure 5.10: Pre-fit data–background comparison of the m_{HH} distributions of in the SR prior to m_{HH} selection. The uncertainty includes statistical and experimental uncertainties of the background, shown in the hatched band. Two $X \to HH$ signals are overlaid in dashed lines, normalising to their cross section at corresponding upper limits. The first and last bins contain the underflow and overflow contents.

event yields in Table 5.8. The detailed numbers are summarised in Table 5.9. The discontinuities of the upper limits at mass hypotheses equals 1.6 and 2.5 TeV are due to different selections that are applied on m_{HH} (see Section 5.2.2).

The relative impact of different sources of systematic uncertainties on the upper limits is measured by let in or out a particular group of systematic uncertainties in the upper limit calculation. At mass hypotheses lower than 1.2 TeV, the largest impact is from the uncertainties in large-R jet energy and mass scales. The total impact for 1 TeV mass point is 85%. For mass hypotheses from 1.2 to 2.0 TeV, the total impact is about 20%, dominated by the uncertainties in di-τ tagging. For 2.5 TeV and beyond, the total impact is less than 10%.

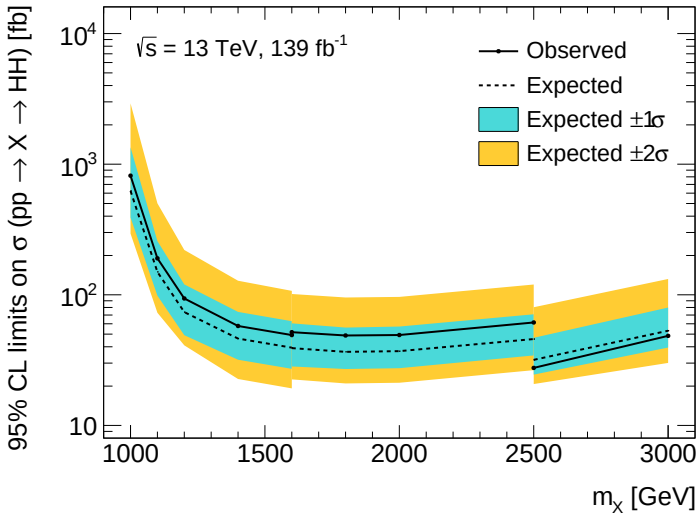


Figure 5.11: Observed (solid line) and expected (dashed line) upper limits on the cross section of $gg \to X \to HH$ at 95% CL in the search for resonant Higgs boson pair production via a narrow-width scalar particle (X) as a function of m_X in the boosted analysis. The markers indicate that the upper limits are calculated based on the signal templates generated by simulated samples with the corresponding mass. The discontinuities of the upper limits at mass hypotheses equals 1.6 and 2.5 TeV are due to different selections that are applied on m_{HH}. For the mass hypotheses with the same m_{HH} selection, the intermediate upper limits are linearly interpolated. The green and yellow bands stand for the $\pm 1, 2$-σ variations on the expected limits, respectively.

Table 5.9: Summary of the observed and expected upper limits, as well as the $\pm 1, 2\text{-}\sigma$ variations of the expected upper limits on the cross section of $gg \to X \to HH$ at 95% CL in the search for resonant Higgs boson pair production via a narrow-width scalar particle (X) of various m_X. The upper limits are calculated in the SR with different requirements on m_{HH}, as stated in the table.

m_X [TeV]	95% CL upper limits on $\sigma_{gg\to X\to HH}$ [fb]					
	Observed	-2σ	-1σ	Expected	1σ	2σ
$m_{HH} > 0$ TeV						
1.0	820	294	390	624	1300	2900
1.1	190	73	99	150	260	500
1.2	94	41	49	74	120	220
1.4	58	23	32	46	74	130
1.6	49	19	27	39	63	110
$m_{HH} > 0.9$ TeV						
1.6	52	23	28	39	60	100
1.8	49	21	27	37	56	95
2.0	49	21	27	37	57	96
2.5	61	26	34	46	71	120
$m_{HH} > 1.2$ TeV						
2.5	28	21	24	32	47	80
3.0	49	30	39	53	80	130

Chapter 6

Conclusion and Outlook

A search for non-resonant and resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel is carried out, using an integrated luminosity of 139 fb^{-1} proton-proton collision data collected by the ATLAS experiment during the Run-2 era of the LHC. Two analyses targetting at different topologies of final-state physics objects and different production mode of HH are presented — the resolved analysis and the boosted analysis. The $\tau_{\text{had}}\tau_{\text{had}}$ decay channel is studied at length. This channel is statistically combined with the $\tau_{\text{lep}}\tau_{\text{had}}$ channel in the resolved analysis.

The resolved analysis improves the expected sensitivity by a factor of four compared with the previous ATLAS search in the same HH decay channel [9]. Half of that is due to the larger dataset, and the other half profits by the advanced reconstruction and identification of b-jets and hadronic τ-leptons, as well as the improved multivariate analysis and background modelling. The boosted analysis extends the sensitive search range of resonance particle mass in the resonant HH production by using a novel hadronic boosted di-τ tagging technique. In the same analysis, the efficiency of di-τ tagging is calibrated.

No significant excesses of events are observed above the expectations of the Standard Model background. The most significant deviation is observed in the search for resonant HH production at a resonance mass of 1 TeV in the resolved analysis, which corresponds to a local (global) significance of 2.8σ (1.7σ) considering only $\tau_{\text{had}}\tau_{\text{had}}$ channel, and the results are 3.0σ (2.0σ) after combining the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. Upper limits on the HH production cross section at 95% confidence-level based on CL_s technique are determined in both the search for non-resonant and resonant HH production.

In the search for non-resonant HH production, the observed (expected) upper limits on the production cross section are 5.0 (4.4) times the SM expectation of ggF+VBF HH production cross section using the $\tau_{\text{had}}\tau_{\text{had}}$ channel. Combined with the $\tau_{\text{lep}}\tau_{\text{had}}$ channel, the corresponding observed (expected) cross section upper limits are 4.7 (3.9) times the ggF+VBF HH production cross section. The $b\bar{b}\tau^+\tau^-$ decay channel has so far the best expected sensitivity compared with other HH decay channels in ATLAS [176].

The same analysis strategy as that in the search for non-resonant HH production is utilised to constrain the strength of Higgs boson self-coupling. Only the ggF HH production mode is

considered. The values of self-coupling modifier κ_λ that lie outside the observed (expected) ranges of $[-2.5, 9.7]$ ($[-2.3, 9.8]$) are excluded at 95% confidence-level. Recently, the results are updated with the inclusion of the VBF production mode and the combination with the $\tau_{\text{lep}}\tau_{\text{had}}$ channel in Ref. [132], where the values of κ_λ are constrained to the observed (expected) ranges of $[-2.4, 9.2]$ ($[-2.0, 9.0]$).

In the search for resonant HH production, the upper limits on the HH production cross section via the decay of a heavy, narrow-width, scalar resonance particle are determined. In the resolved analysis, for resonance particle mass between 251 and 1600 GeV，the observed (expected) upper limits are 27–1600 fb (18–1000 fb) considering only the $\tau_{\text{had}}\tau_{\text{had}}$ channel and 23–930 fb (12–840 fb) after combining the $\tau_{\text{lep}}\tau_{\text{had}}$ channel. In the boosted analysis where only the $\tau_{\text{had}}\tau_{\text{had}}$ channel is exploited, the observed (expected) upper limits are 28–94 fb (32–74 fb) for resonance particle mass between 1.2 and 3.0 TeV. The exact value depends on the mass hypothesis of the resonance particle. The upper limits are equally or more stringent compared with other HH decay channels in most of the mass search range [176].

The next step in the search for non-resonant HH production mode is to optimise the analysis strategy to constrain the strength of Higgs boson self-coupling. This is not yet done in the analysis presented in this thesis (Section 4.8), in which the analysis strategy is optimised for the SM-like HH kinematics, i.e., a single point of $\kappa_\lambda = 1$. Meanwhile, the constraints on κ_λ determined by the precision measurement on single Higgs boson production can be combined with those determined by the search for Higgs boson pair production, as reported in Ref. [177] using part of the Run-2 data. Searching for the quartic coupling of Higgs bosons and vector bosons (see Figure 1.3) is another interesting topic about the non-resonant HH production mode. The VBF HH production mode uniquely provide a way to directly measure the strength of this coupling [178–180]. Regarding the search for resonant HH production, there is much room for improvements in the boosted analysis. The signal acceptance in the boosted analysis is mainly limited by the p_{T} thresholds of the large-R jet triggers and the offline p_{T} selections, as well as the efficiency of di-τ reconstruction and identification. The choice of triggers may be revisited to gain more signal acceptance, and the di-τ tagging algorithms can be tuned to optimise the sensitivity. It is also expected that the improved b-tagging method [79, 80] can further improve the sensitivity.

Improved performance in the reconstruction and identification of physics objects are foreseen in the upcoming Run-3 data-taking. In this thesis, a new DeepSet τ_{had} decay mode classification method is developed. The implementation of this method has been integrated into the ATLAS software for Run-3 physics analysis. The method significantly improves the classification efficiency and purity for all τ_{had} decay modes in most regions of interest (e.g. τ_{had} with low p_{T}), compared with the previous PanTau method. The relative improvement in both classification efficiency and purity is about 9%. The application of this new method in future physics analyses is promising, for instance, in the analysis of the CP structure of the Yukawa coupling between the Higgs boson and τ leptons.

Looking ahead to future studies, we expect to see evidence of Higgs boson pair production via SM processes in the era of the high-luminosity LHC. The sensitivity of the $b\bar{b}\tau^+\tau^-$ decay channel in the high-luminosity LHC has been projected based on the combined results of $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels in this thesis [181]. The expected significance of observing an excess of SM HH events over the SM background events is 2.8σ [1]. Combining with the $b\bar{b}\gamma\gamma$ decay channel of HH, the expected significance reaches 3.2σ [182]. While we are excited about the possible observation of the SM HH production, we would also wish to see excesses or anomalies originating from resonant HH production or non-SM couplings that would break through the dominance of the SM.

[1] These numbers are derived by assuming the b-tagging and theoretical uncertainties are halved, and the τ_{had} uncertainties that originate from data statistics are ignored.

Bibliography

[1] ATLAS Collaboration, *Observation of a new particle in the search for the Standard Model Higgs boson with the ATLAS detector at the LHC*, Phys. Lett. B **716** (2012) 1.

[2] CMS Collaboration, *Observation of a new boson at a mass of 125 GeV with the CMS experiment at the LHC*, Phys. Lett. B **716** (2012) 30.

[3] ATLAS Collaboration, *Combined measurements of Higgs boson production and decay using up to* 139 fb^{-1} *of proton-proton collision data at* $\sqrt{s} = 13$ *TeV collected with the ATLAS experiment*, ATLAS-CONF-2021-053, 2021.

[4] CMS Collaboration, *Combined Higgs boson production and decay measurements with up to 137 fb-1 of proton-proton collision data at sqrts = 13 TeV*, CMS-PAS-HIG-19-005, 2020.

[5] D. de Florian et al., *Handbook of LHC Higgs Cross Sections: 4. Deciphering the Nature of the Higgs Sector*, **2/2017** (2016).

[6] J. Alison et al., *Higgs boson potential at colliders: Status and perspectives*, Rev. Phys. **5** (2020) 100045, ed. by Biagio Di Micco, Maxime Gouzevitch, Javier Mazzitelli, and Caterina Vernieri.

[7] ATLAS Collaboration, *Combination of searches for Higgs boson pairs in* pp *collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Lett. B **800** (2020) 135103.

[8] CMS Collaboration, *Combination of Searches for Higgs Boson Pair Production in Proton–Proton Collisions at* $\sqrt{s} = 13$ *TeV*, Phys. Rev. Lett. **122** (2019) 121803.

[9] ATLAS Collaboration, *A search for resonant and non-resonant Higgs boson pair production in the* $b\bar{b}\tau^+\tau^-$ *decay channel in* pp *collisions at* $\sqrt{s} = 13$ *TeV with the ATLAS detector*, Phys. Rev. Lett. **121** (2018) 191801, Erratum: Phys. Rev. Lett. **122** (2019) 089901.

[10] CMS Collaboration, *Search for Higgs boson pair production in the* $bb\tau\tau$ *final state in proton–proton collisions at* $\sqrt{s} = 8$ *TeV*, Phys. Rev. D **96** (2017) 072004.

[11] Mark Thomson, *Modern Particle Physics*, Cambridge University Press, 2013.

[12] Francis Halzen and Alan D. Martin, *Quarks and Leptons: An Introductory Course in Modern Particle Physics*, 1984.

[13] John F. Gunion, *The Higgs Hunter's Guide*, 1990.

[14] Sheldon L. Glashow, *Partial-symmetries of weak interactions*, Nuclear Physics **22** (1961) 579.

[15] A. Salam and J.C. Ward, *Electromagnetic and weak interactions*, Physics Letters **13** (1964) 168.

[16] Steven Weinberg, *A Model of Leptons*, Phys. Rev. Lett. **19** (21 1967) 1264.

[17] Peter W. Higgs, *Broken symmetries, massless particles and gauge fields*, Physics Letters **12** (1964) 132.

[18] Peter W. Higgs, *Broken Symmetries and the Masses of Gauge Bosons*, Phys. Rev. Lett. **13** (16 1964) 508.

[19] Peter W. Higgs, *Spontaneous Symmetry Breakdown without Massless Bosons*, Phys. Rev. **145** (4 1966) 1156.

[20] F. Englert and R. Brout, *Broken Symmetry and the Mass of Gauge Vector Mesons*, Phys. Rev. Lett. **13** (9 1964) 321.

[21] G. S. Guralnik, C. R. Hagen, and T. W. B. Kibble, *Global Conservation Laws and Massless Particles*, Phys. Rev. Lett. **13** (20 1964) 585.

[22] T. W. B. Kibble, *Symmetry Breaking in Non-Abelian Gauge Theories*, Phys. Rev. **155** (5 1967) 1554.

[23] Particle Data Group, *Review of Particle Physics*, Progress of Theoretical and Experimental Physics **2020** (2020), 083C01.

[24] R. Frederix et al., *Higgs pair production at the LHC with NLO and parton-shower effects*, Phys. Lett. B **732** (2014) 142.

[25] J. Baglio et al., $gg \to HH$ *: Combined uncertainties*, Phys. Rev. D **103** (2021) 056002.

[26] Frédéric A. Dreyer and Alexander Karlberg, *Vector-Boson Fusion Higgs Pair Production at* N^3LO, Phys. Rev. D **98** (2018) 114016.

[27] Simone Alioli, Paolo Nason, Carlo Oleari, and Emanuele Re, *A general framework for implementing NLO calculations in shower Monte Carlo programs: the POWHEG BOX*, JHEP **06** (2010) 043.

[28] S. Borowka et al., *Higgs Boson Pair Production in Gluon Fusion at Next-to-Leading Order with Full Top-Quark Mass Dependence*, Phys. Rev. Lett. **117** (2016) 012001, [Erratum: Phys.Rev.Lett. 117, 079901 (2016)].

[29] J. Alwall et al., *The automated computation of tree-level and next-to-leading order differential cross sections, and their matching to parton shower simulations*, JHEP **07** (2014) 079.

[30] Rikkert Frederix and Stefano Frixione, *Merging meets matching in MC@NLO*, JHEP **12** (2012) 061.

[31] ATLAS Collaboration, *Validation of signal Monte Carlo event generation in searches for Higgs boson pairs with the ATLAS detector*, ATL-PHYS-PUB-2019-007, 2019.

[32] Dario Buttazzo et al., *Investigating the near-criticality of the Higgs boson*, JHEP **12** (2013) 089.

[33] Giuseppe Degrassi et al., *Higgs mass and vacuum stability in the Standard Model at NNLO*, JHEP **08** (2012) 098.

[34] So Chigusa, Takeo Moroi, and Yutaro Shoji, *State-of-the-Art Calculation of the Decay Rate of Electroweak Vacuum in the Standard Model*, Phys. Rev. Lett. **119** (2017) 211801.

[35] David E. Morrissey and Michael J. Ramsey-Musolf, *Electroweak baryogenesis*, New J. Phys. **14** (2012) 125003.

[36] A. I. Bochkarev and M. E. Shaposhnikov, *Electroweak Production of Baryon Asymmetry and Upper Bounds on the Higgs and Top Masses*, Mod. Phys. Lett. A **2** (1987) 417.

[37] K. Kajantie, M. Laine, K. Rummukainen, and M. Shaposhnikov, *Generic rules for high temperature dimensional reduction and their application to the standard model*, Nuclear Physics B **458** (1996) 90.

[38] Manuel Reichert et al., *Probing baryogenesis through the Higgs boson self-coupling*, Phys. Rev. D **97** (2018) 075008.

[39] Hooman Davoudiasl, Ryuichiro Kitano, Tianjun Li, and Hitoshi Murayama, *The New minimal standard model*, Phys. Lett. B **609** (2005) 117.

[40] Vernon Barger, Paul Langacker, Mathew McCaskey, Michael J. Ramsey-Musolf, and Gabe Shaughnessy, *LHC Phenomenology of an Extended Standard Model with a Real Scalar Singlet*, Phys. Rev. D **77** (2008) 035005.

[41] Jose M. No and Michael Ramsey-Musolf, *Probing the Higgs Portal at the LHC Through Resonant di-Higgs Production*, Phys. Rev. D **89** (2014) 095031.

[42] Alexandra Carvalho et al., *Analytical parametrization and shape classification of anomalous HH production in the EFT approach*, (2016).

[43] G. Buchalla, M. Capozi, A. Celis, G. Heinrich, and L. Scyboz, *Higgs boson pair production in non-linear Effective Field Theory with full m_t-dependence at NLO QCD*, JHEP **09** (2018) 057.

[44] G. Heinrich, S. P. Jones, M. Kerner, G. Luisoni, and L. Scyboz, *Probing the trilinear Higgs boson coupling in di-Higgs production at NLO QCD including parton shower effects*, JHEP **06** (2019) 066.

[45] Lyndon Evans and Philip Bryant, *LHC Machine*, Journal of Instrumentation **3** (2008) S08001.

[46] ATLAS Collaboration, *The ATLAS Experiment at the CERN Large Hadron Collider*, JINST **3** (2008) S08003.

[47] Oliver Sim Brüning et al., *LHC Design Report v.1 the LHC Main Ring*, CERN Yellow Reports: Monographs (2004).

[48] Oliver Sim Brüning et al., *LHC Design Report v.2 the LHC Infrastructure and General Services*, CERN Yellow Reports: Monographs (2004).

[49] Michael Benedikt, Paul Collier, V Mertens, John Poole, and Karlheinz Schindl, *LHC Design Report v.3 the LHC Injector Chain*, CERN Yellow Reports: Monographs (2004).

[50] ATLAS Collaboration, *ATLAS: Technical Proposal for a General-Purpose pp Experiment at the Large Hadron Collider at CERN*, 1994.

[51] CMS Collaboration, *CMS, the Compact Muon Solenoid: technical proposal*, LHC technical proposal, Geneva: CERN, 1994.

[52] LHCb Collaboration, *LHCb : Technical Proposal*, Geneva: CERN, 1998.

[53] ALICE Collaboration, *ALICE: Technical proposal for a Large Ion collider Experiment at the CERN LHC*, LHC technical proposal, Geneva: CERN, 1995.

[54] Julie Haffner, *Complexe des accélérateurs du CERN*, `http://cds.cern.ch/record/1621894`, General Photo, 2013.

[55] ATLAS Collaboration, *Public results of Run2 Luminosity*, `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/LuminosityPublicResultsRun2`.

[56] O. Aberle et al., *High-Luminosity Large Hadron Collider (HL-LHC): Technical design report*, CERN Yellow Reports: Monographs, Geneva: CERN, 2020.

[57] CERN, *Dessin representant le detecteur ATLAS*, `https://cds.cern.ch/record/39038`, 1998.

[58] M Capeans et al., *ATLAS Insertable B-Layer Technical Design Report*, 2010.

[59] ATLAS Collaboration, *ATLAS liquid-argon calorimeter: Technical Design Report*, Technical design report. ATLAS, Geneva: CERN, 1996.

[60] Georges Aad et al., *Operation of the ATLAS trigger system in Run 2*, JINST **15** (2020) P10004.

[61] ATLAS Collaboration, *Performance of the ATLAS trigger system in 2015*, Eur. Phys. J. C **77** (2017) 317.

[62] Rosa Simoniello, *The ATLAS Level-1 Topological Processor: from design to routine usage in Run-2*, (2018).

[63] ATLAS Collaboration, *Performance of the ATLAS track reconstruction algorithms in dense environments in LHC Run 2*, Eur. Phys. J. C **77** (2017) 673.

[64] ATLAS Collaboration, *Early Inner Detector Tracking Performance in the 2015 Data at* $\sqrt{s} = 13$ *TeV*, ATL-PHYS-PUB-2015-051, 2015.

[65] ATLAS Collaboration, *Electron reconstruction and identification in the ATLAS experiment using the 2015 and 2016 LHC proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **79** (2019) 639.

[66] ATLAS Collaboration, *Electron and photon energy calibration with the ATLAS detector using 2015–2016 LHC proton–proton collision data*, JINST **14** (2019) P03017.

[67] ATLAS Collaboration, *Muon reconstruction performance of the ATLAS detector in proton–proton collision data at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **76** (2016) 292.

[68] ATLAS Collaboration, *Muon reconstruction and identification efficiency in ATLAS using the full Run 2 pp collision data set at* $\sqrt{s} = 13$ *TeV*, (2020).

[69] ATLAS Collaboration, *Jet reconstruction and performance using particle flow with the ATLAS Detector*, Eur. Phys. J. C **77** (2017) 466.

[70] ATLAS Collaboration, *Identification of boosted Higgs bosons decaying into b-quark pairs with the ATLAS detector at* 13 *TeV*, Eur. Phys. J. C **79** (2019) 836.

[71] ATLAS Collaboration, *Performance of missing transverse momentum reconstruction with the ATLAS detector using proton–proton collisions at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **78** (2018) 903.

[72] T Cornelissen et al., *The new ATLAS track reconstruction (NEWT)*, Journal of Physics: Conference Series **119** (2008) 032014.

[73] R. Frühwirth, *Application of Kalman filtering to track and vertex fitting*, Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment **262** (1987) 444.

[74] ATLAS Collaboration, *Improved electron reconstruction in ATLAS using the Gaussian Sum Filter-based model for bremsstrahlung*, ATLAS-CONF-2012-047, 2012.

[75] Matteo Cacciari, Gavin P. Salam, and Gregory Soyez, *The anti-k_t jet clustering algorithm*, JHEP **04** (2008) 063.

[76] ATLAS Collaboration, *Topological cell clustering in the ATLAS calorimeters and its performance in LHC Run 1*, Eur. Phys. J. C **77** (2017) 490.

[77] ATLAS Collaboration, *Selection of jets produced in 13TeV proton-proton collisions with the ATLAS detector*, (2015).

[78] ATLAS Collaboration, *Tagging and suppression of pileup jets*, (2014).

[79] ATLAS Collaboration, *ATLAS b-jet identification performance and efficiency measurement with* $t\bar{t}$ *events in pp collisions at* $\sqrt{s} = 13$ *TeV*, Eur. Phys. J. C **79** (2019) 970.

[80] ATLAS Collaboration, *Optimisation and performance studies of the ATLAS b-tagging algorithms for the 2017-18 LHC run*, (2017).

[81] ATLAS Collaboration, *Evidence for the $H \to b\bar{b}$ decay with the ATLAS detector*, JHEP **12** (2017) 024.

[82] A.D.Bukin, *Fitting function for asymmetric peaks*, (2007).

[83] David Krohn, Jesse Thaler, and Lian-Tao Wang, *Jet trimming*, Journal of High Energy Physics **2010** (2010) 84.

[84] S. Catani, Yu.L. Dokshitzer, M.H. Seymour, and B.R. Webber, *Longitudinally invariant $K(t)$ clustering algorithms for hadron-hadron collisions*, Nuclear Physics B **406** (1993) 187.

[85] Stephen D. Ellis and Davison E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. **D48** (1993) 3160.

[86] Stephen D. Ellis and Davison E. Soper, *Successive combination jet algorithm for hadron collisions*, Phys. Rev. D **48** (7 1993) 3160.

[87] ATLAS Collaboration, *Identification of boosted, hadronically decaying W bosons and comparisons with ATLAS data taken at $\sqrt{s} = 8$ TeV*, Eur. Phys. J. C **76** (2016) 154.

[88] ATLAS Collaboration, *Jet mass reconstruction with the ATLAS Detector in early Run 2 data*, ATLAS-CONF-2016-035, 2016.

[89] David Krohn, Jesse Thaler, and Lian-Tao Wang, *Jets with variable R*, Journal of High Energy Physics **2009** (2009) 059.

[90] ATLAS Collaboration, *Variable Radius, Exclusive-k_T, and Center-of-Mass Subjet Reconstruction for Higgs($\to b\bar{b}$) Tagging in ATLAS*, ATL-PHYS-PUB-2017-010, 2017.

[91] Matteo Cacciari and Gavin P. Salam, *Pileup subtraction using jet areas*, Physics Letters B **659** (2008) 119.

[92] Matteo Cacciari, Gavin P Salam, and Gregory Soyez, *The catchment area of jets*, Journal of High Energy Physics **2008** (2008) 005.

[93] ATLAS Collaboration, *Optimisation of the ATLAS b-tagging performance for the 2016 LHC Run*, ATL-PHYS-PUB-2016-012, 2016.

[94] ATLAS Collaboration, *ATLAS data quality operations and performance for 2015–2018 data-taking*, JINST **15** (2020) P04003.

[95] ATLAS Collaboration, *The ATLAS Simulation Infrastructure*, Eur. Phys. J. C **70** (2010) 823.

[96] G. Altarelli and G. Parisi, *Asymptotic freedom in parton language*, Nuclear Physics B **126** (1977) 298.

[97] GEANT4 Collaboration, S. Agostinelli, et al., *GEANT4 – a simulation toolkit*, Nucl. Instrum. Meth. A **506** (2003) 250.

[98] M Beckingham et al., *The simulation principle and performance of the ATLAS fast calorimeter simulation FastCaloSim*, (2010).

[99] ATLAS Collaboration, *Jet energy scale and resolution measured in proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, (2020).

[100] ATLAS Collaboration, *In situ calibration of large-radius jet energy and mass in 13 TeV proton–proton collisions with the ATLAS detector*, Eur. Phys. J. C **79** (2019) 135.

[101] ATLAS Collaboration, *Measurement of the tau lepton reconstruction and identification performance in the ATLAS experiment using pp collisions at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2017-029, 2017.

[102] ATLAS Collaboration, *Athena*, `https://doi.org/10.5281/zenodo.2641997`, 2019.

[103] Pushpalatha C. Bhat, *Multivariate Analysis Methods in Particle Physics*, Annual Review of Nuclear and Particle Science **61** (2011) 281.

[104] Alexander Radovic et al., *Machine learning at the energy and intensity frontiers of particle physics*, Nature **560** (2018) 41.

[105] Dimitri Bourilkov, *Machine and Deep Learning Applications in Particle Physics*, Int. J. Mod. Phys. A **34** (2020) 1930019.

[106] Dan Guest, Kyle Cranmer, and Daniel Whiteson, *Deep Learning and its Application to LHC Physics*, Ann. Rev. Nucl. Part. Sci. **68** (2018) 161.

[107] Christopher Bishop, *Pattern Recognition and Machine Learning*, vol. 16, 2006 140.

[108] L. Breiman, Jerome H. Friedman, Richard A. Olshen, and C. J. Stone, "Classification and Regression Trees", 1983.

[109] Yoav Freund and Robert E Schapire, *A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting*, Journal of Computer and System Sciences **55** (1997) 119.

[110] C. Van Der Malsburg, "Frank Rosenblatt: Principles of Neurodynamics: Perceptrons and the Theory of Brain Mechanisms", *Brain Theory*, ed. by Günther Palm and Ad Aertsen, Berlin, Heidelberg: Springer Berlin Heidelberg, 1986 245.

[111] Russell Reed and Robert J Marks, *Neural Smithing: Supervised Learning in Feedforward Artificial Neural Networks*, The MIT Press, 1999.

[112] Andreas Hoecker et al., *TMVA - Toolkit for Multivariate Data Analysis*, 2007.

[113] ATLAS Collaboration, *Reconstruction, Energy Calibration, and Identification of Hadronically Decaying Tau Leptons in the ATLAS Experiment for Run-2 of the LHC*, ATL-PHYS-PUB-2015-045, 2015.

[114] Dirk Uwe Duschinger, "Search for neutral bosons decaying into the fully hadronic di-tau final state with the ATLAS detector at the LHC", Presented 07 Feb 2019, 2018.

[115] ATLAS Collaboration, *Reconstruction of hadronic decay products of tau leptons with the ATLAS experiment*, Eur. Phys. J. C **76** (2016) 295.

[116] ATLAS Collaboration, *Identification of hadronic tau lepton decays using neural networks in the ATLAS experiment*, ATL-PHYS-PUB-2019-033, 2019.

[117] ATLAS Collaboration, *Identification and energy calibration of hadronically decaying tau leptons with the ATLAS experiment in pp collisions at* $\sqrt{s} = 8$ *TeV*, Eur. Phys. J. C **75** (2015) 303.

[118] ATLAS Collaboration, *Probing the CP nature of the Higgs boson coupling to* τ *leptons at HL-LHC*, ATL-PHYS-PUB-2019-008, 2019.

[119] CMS Collaboration, *Analysis of the CP structure of the Yukawa coupling between the Higgs boson and* τ *leptons in proton-proton collisions at* $\sqrt{s}$ *= 13 TeV*, (2021).

[120] Michael M. Bronstein, Joan Bruna, Yann LeCun, Arthur Szlam, and Pierre Vandergheynst, *Geometric deep learning: going beyond Euclidean data*, CoRR **abs/1611.08097** (2016).

[121] Manzil Zaheer et al., *Deep Sets*, CoRR **abs/1703.06114** (2017).

[122] ATLAS Collaboration, *ATLAS Pythia 8 tunes to* 7 *TeV data*, ATL-PHYS-PUB-2014-021, 2014.

[123] Richard D. Ball et al., *Parton distributions with LHC data*, Nucl. Phys. B **867** (2013) 244.

[124] ATLAS Collaboration, *Characterization of Interaction-Point Beam Parameters Using the pp Event-Vertex Distribution Reconstructed in the ATLAS Detector at the LHC*, ATLAS-CONF-2010-027, 2010.

[125] T Barillari et al., *Local Hadronic Calibration*, tech. rep., CERN, 2008.

[126] Benedict Tobias Winter, "Measurement of Tau Polarisation in $Z/\gamma^* \to \tau\tau$ Decays in Proton−Proton Collisions at ATLAS", PhD thesis: U. Bonn (main), 2019.

[127] Jonathan Shlomi, Peter Battaglia, and Jean-Roch Vlimant, *Graph Neural Networks in Particle Physics*, (2020).

[128] ATLAS Collaboration, *Deep Sets based Neural Networks for Impact Parameter Flavour Tagging in ATLAS*, ATL-PHYS-PUB-2020-014, 2020.

[129] François Chollet et al., *Keras*, `https://keras.io`, 2015.

[130] ATLAS Collaboration, *Reconstruction and identification of boosted di-*τ *systems in a search for Higgs boson pairs using* 13 *TeV proton–proton collision data in ATLAS*, JHEP **11** (2020) 163.

[131] ATLAS Collaboration, *Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 TeV pp collision data from the ATLAS detector*, ATLAS-CONF-2021-030, 2021.

[132] ATLAS Collaboration, *Combination of searches for non-resonant and resonant Higgs boson pair production in the $b\bar{b}\gamma\gamma$, $b\bar{b}\tau^+\tau^-$ and $b\bar{b}b\bar{b}$ decay channels using pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, ATLAS-CONF-2021-052, 2021.

[133] Jon Butterworth et al., *PDF4LHC recommendations for LHC Run II*, J. Phys. G **43** (2016) 023001.

[134] Richard D. Ball et al., *Parton distributions for the LHC run II*, JHEP **04** (2015) 040.

[135] Torbjörn Sjöstrand et al., *An introduction to PYTHIA 8.2*, Comput. Phys. Commun. **191** (2015) 159.

[136] Johannes Bellm et al., *Herwig 7.1 Release Note*, (2017).

[137] Enrico Bothmann et al., *Event generation with Sherpa 2.2*, SciPost Phys. **7** (2019) 034.

[138] ATLAS Collaboration, *The ATLAS Tau Trigger in Run 2*, ATLAS-CONF-2017-061, 2017.

[139] ATLAS Collaboration, *Public Tau Trigger Plots for Collision Data*, `https://twiki.cern.ch/twiki/bin/view/AtlasPublic/TauTriggerPublicResults`.

[140] A. Elagin, P. Murat, A. Pranko, and A. Safonov, *A New Mass Reconstruction Technique for Resonances Decaying to di-tau*, Nucl. Instrum. Meth. A **654** (2011) 481.

[141] Pierre Baldi, Kyle Cranmer, Taylor Faucett, Peter Sadowski, and Daniel Whiteson, *Parameterized neural networks for high-energy physics*, Eur. Phys. J. C **76** (2016) 235.

[142] Glen Cowan, Kyle Cranmer, Eilam Gross, and Ofer Vitells, *Asymptotic formulae for likelihood-based tests of new physics*, Eur. Phys. J. C **71** (2011) 1554, Erratum: Eur. Phys. J. C **73** (2013) 2501.

[143] ATLAS Collaboration, *ATLAS simulation of boson plus jets processes in Run 2*, ATL-PHYS-PUB-2017-006, 2017.

[144] ATLAS Collaboration, *Performance of the ATLAS muon triggers in Run 2*, JINST **15** (2020) P09015.

[145] ATLAS Collaboration, *Performance of electron and photon triggers in ATLAS during LHC Run 2*, Eur. Phys. J. C **80** (2020) 47.

[146] ATLAS Collaboration, *Measurements of top-quark pair differential cross-sections in the lepton+jets channel in pp collisions at $\sqrt{s}$ = 13 TeV using the ATLAS detector*, JHEP **11** (2017) 191.

[147] ATLAS Collaboration, *Measurement of the $t\bar{t}$ production cross-section in the lepton+jets channel at $\sqrt{s}$ = 13 TeV with the ATLAS experiment*, Phys. Lett. B **810** (2020) 135797.

[148] Michal Czakon et al., *Top-pair production at the LHC through NNLO QCD and NLO EW*, JHEP **10** (2017) 186.

[149] ATLAS Collaboration, *Measurements of top-quark pair single- and double-differential cross-sections in the all-hadronic channel in pp collisions at $\sqrt{s} = 13$ TeV using the ATLAS detector*, JHEP **01** (2021) 033.

[150] G. Avoni et al., *The new LUCID-2 detector for luminosity measurement and monitoring in ATLAS*, Journal of Instrumentation **13** (2018) P07017.

[151] ATLAS Collaboration, *Electron and photon performance measurements with the ATLAS detector using the 2015–2017 LHC proton–proton collision data*, JINST **14** (2019) P12006.

[152] ATLAS Collaboration, *Performance of pile-up mitigation techniques for jets in pp collisions at $\sqrt{s} = 8$ TeV using the ATLAS detector*, Eur. Phys. J. C **76** (2016) 581.

[153] ATLAS Collaboration, *Measurement of b-tagging efficiency of c-jets in $t\bar{t}$ events using a likelihood approach with the ATLAS detector*, ATLAS-CONF-2018-001, 2018.

[154] ATLAS Collaboration, *Calibration of light-flavour b-jet mistagging rates using ATLAS proton–proton collision data at $\sqrt{s} = 13$ TeV*, ATLAS-CONF-2018-006, 2018.

[155] John Kenneth Anders, "Searches for direct pair production of third generation squarks, and dark matter, in final states containing $b-$jets and $E_{\mathrm{T}}^{\mathrm{miss}}$ using the ATLAS detector at the LHC", Presented 07 Jun 2017, 2017.

[156] S. Catani, F. Krauss, R. Kuhn, and B. R. Webber, *QCD Matrix Elements + Parton Showers*, JHEP **11** (2001) 063.

[157] ATLAS Collaboration, *Study of top-quark pair modelling and uncertainties using ATLAS measurements at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2020-023, 2020.

[158] ATLAS Collaboration, *Improvements in $t\bar{t}$ modelling using NLO+PS Monte Carlo generators for Run 2*, ATL-PHYS-PUB-2018-009, 2018.

[159] ATLAS Collaboration, *Simulation of top-quark production for the ATLAS experiment at $\sqrt{s} = 13$ TeV*, ATL-PHYS-PUB-2016-004, 2016.

[160] Stefano Frixione, Eric Laenen, Patrick Motylinski, Bryan R. Webber, and Chris D. White, *Single-top hadroproduction in association with a W boson*, JHEP **07** (2008) 029.

[161] ATLAS Collaboration, *Study of heavy flavor quarks produced in association with top quark pairs at $\sqrt{s} = 7$ TeV using the ATLAS detector*, Phys. Rev. D **89** (2014) 072012.

[162] ATLAS Collaboration, *Measurement of the cross-section for W boson production in association with b-jets in pp collisions at $\sqrt{s} = 7$ TeV with the ATLAS detector*, JHEP **06** (2013) 084.

[163] ATLAS Collaboration, *Measurement of VH, $H \to b\bar{b}$ production as a function of the vector-boson transverse momentum in 13 TeV pp collisions with the ATLAS detector*, JHEP **05** (2019) 141.

[164] Suhasini Subba Rao, *Advanced statistical inference*, Texas A & M University, Tech. Rep (2017).

[165] Roger Barlow and Christine Beeston, *Fitting using finite Monte Carlo samples*, Computer Physics Communications **77** (1993) 219.

[166] Kyle Cranmer, George Lewis, Lorenzo Moneta, Akira Shibata, and Wouter Verkerke, *HistFactory: A tool for creating statistical models for use with RooFit and RooStats*, tech. rep., New York U., 2012.

[167] F. James and M. Roos, *Minuit: A System for Function Minimization and Analysis of the Parameter Errors and Correlations*, Comput. Phys. Commun. **10** (1975) 343.

[168] Louis Lyons, *Open statistical issues in Particle Physics*, The Annals of Applied Statistics **2** (2008) 887.

[169] Alexander L. Read, *Presentation of search results: the CL_S technique*, J. Phys. G **28** (2002) 2693.

[170] S. S. Wilks, *The Large-Sample Distribution of the Likelihood Ratio for Testing Composite Hypotheses*, Annals Math. Statist. **9** (1938) 60.

[171] Abraham Wald, *Tests of Statistical Hypotheses Concerning Several Parameters When the Number of Observations is Large*, Transactions of the American Mathematical Society **54** (1943) 426.

[172] Abe Sklar, *Fonctions de repartition a n dimensions et leurs marges*, Publ. Inst. Statist. Univ. Paris **8** (1959) 229.

[173] Thorsten Schmidt, *Coping with copulas*, (2007).

[174] ATLAS Collaboration, *Evaluating statistical uncertainties and correlations using the bootstrap method*, ATL-PHYS-PUB-2021-011, 2021.

[175] ATLAS Collaboration, *Measurement of the associated production of a Higgs boson decaying into b-quarks with a vector boson at high transverse momentum in pp collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, Phys. Lett. B **816** (2021) 136204.

[176] ATLAS Collaboration, *Summary of non-resonant and resonant Higgs boson pair searches from the ATLAS experiment*, ATL-PHYS-PUB-2021-031, 2021.

[177] ATLAS Collaboration, *Constraints on the Higgs boson self-coupling from the combination of single-Higgs and double-Higgs production analyses performed with the ATLAS experiment*, ATLAS-CONF-2019-049, 2019.

[178] ATLAS Collaboration, *Search for the $HH \to b\bar{b}b\bar{b}$ process via vector-boson fusion production using proton–proton collisions at $\sqrt{s} = 13$ TeV with the ATLAS detector*, JHEP **07** (2020) 108.

[179] CMS Collaboration, *Search for nonresonant Higgs boson pair production in final states with two bottom quarks and two photons in proton–proton collisions at $\sqrt{s} = 13$ TeV*, JHEP **03** (2021) 257.

[180] CMS Collaboration, *Search for Higgs boson pair production in the four b quark final state in proton-proton collisions at $\sqrt{s} = 13$ TeV*, (2022).

[181] ATLAS Collaboration, *Projected sensitivity of Higgs boson pair production in the $bb\tau\tau$ final state using proton-proton collisions at HL-LHC with the ATLAS detector*, ATL-PHYS-PUB-2021-044, 2021.

[182] ATLAS Collaboration, *Projected sensitivity of Higgs boson pair production combining the $b\bar{b}\gamma\gamma$ and $b\bar{b}\tau^+\tau^-$ final states with the ATLAS detector at the HL-LHC*, ATL-PHYS-PUB-2022-005, 2022.

[183] Rene Brun and Fons Rademakers, *ROOT – An object oriented data analysis framework*, Nucl. Instrum. Meth. A **389** (1997) 81.

Appendix A

Details on $\tau_{\mathbf{had}}$ decay mode classification

Figure A.1 shows the visualisation of particle flow objects in $\Delta\phi$(object, trackECal)–$\Delta\eta$(object, trackECal) plane.

Figure A.2 summaries the efficiency of identifying the correct decay modes as functions of the generator-level τ_{had} p_{T}, η and average number of bunch crossing. Figure A.3 summaries the mis-identification rate of the decay modes. The former and latter correspond to the diagonal and off-diagonal elements in the efficiency matrix (see Figure 3.8(a)). Figure A.4 summaries the purity of the correctly identified decay modes as functions of the generator-level τ_{had} p_{T}, η and average number of bunch crossing. Figure A.5 summaries the contamination rate of the decay modes. The former and latter correspond to the diagonal and off-diagonal elements in the purity matrix (see Figure 3.8(b)). The figures shown are performed on the DeepSet neural network decay mode classifier with the simulated $\gamma^* \rightarrow \tau\tau$ samples. The vertical error bars indicates the statistical uncertainty of the simulated sample. Figure A.6 compares the signal efficiency and the corresponding 1− background efficiency of the DeepSet and PanTau classifiers in the classifications of $h^{\pm}$ against $h^{\pm}\pi^0$, $h^{\pm}\pi^0$ against $h^{\pm} \geq 2\pi^0$ and $3h^{\pm}$ against $3h^{\pm} \geq 1\pi^0$.

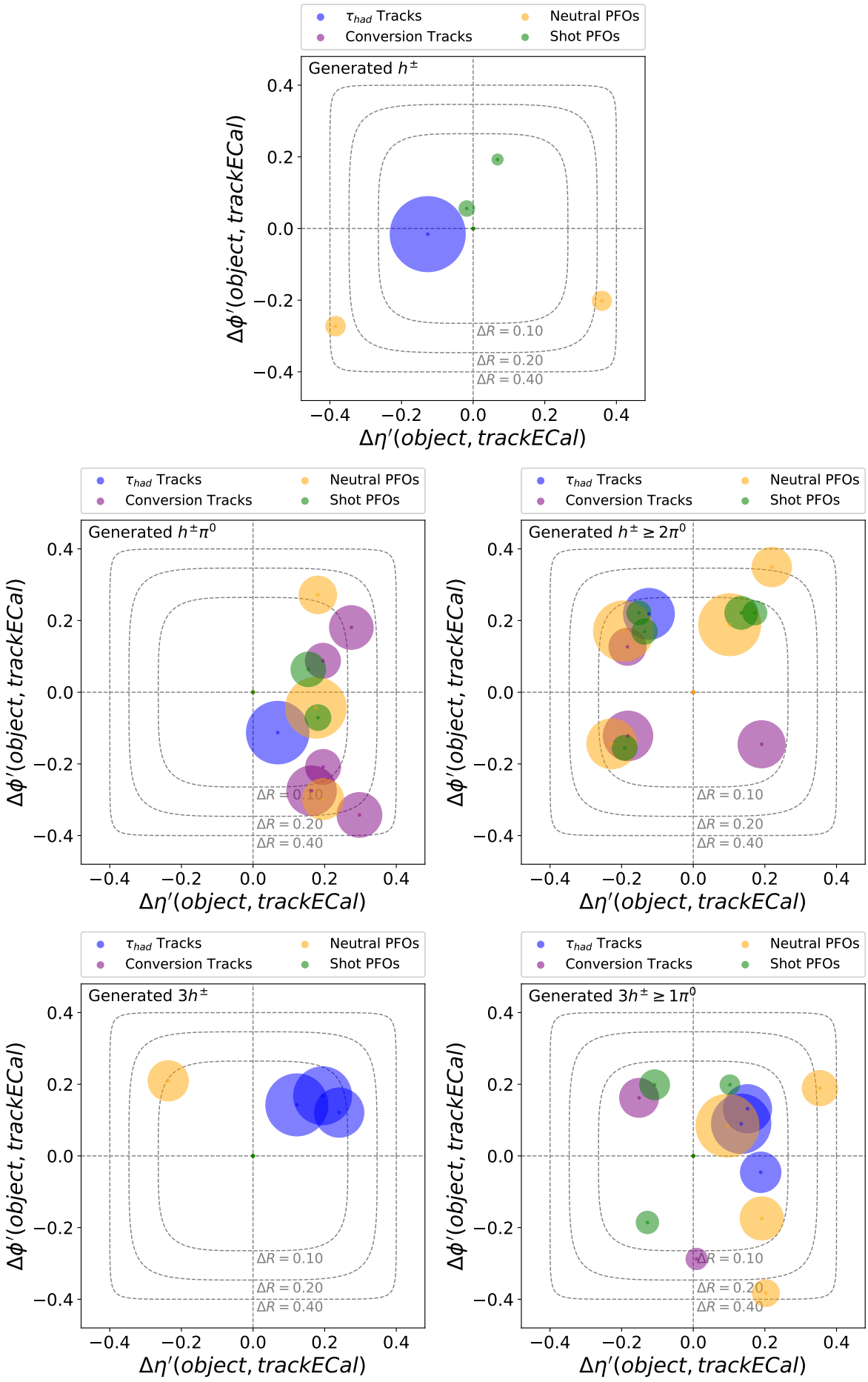


Figure A.1: Visualisation of kinematic variables in displays of various generated τ decay modes with the simulated $\gamma^* \rightarrow \tau\tau$ sample. Details on the figures are explained in Section 3.3.2.1.

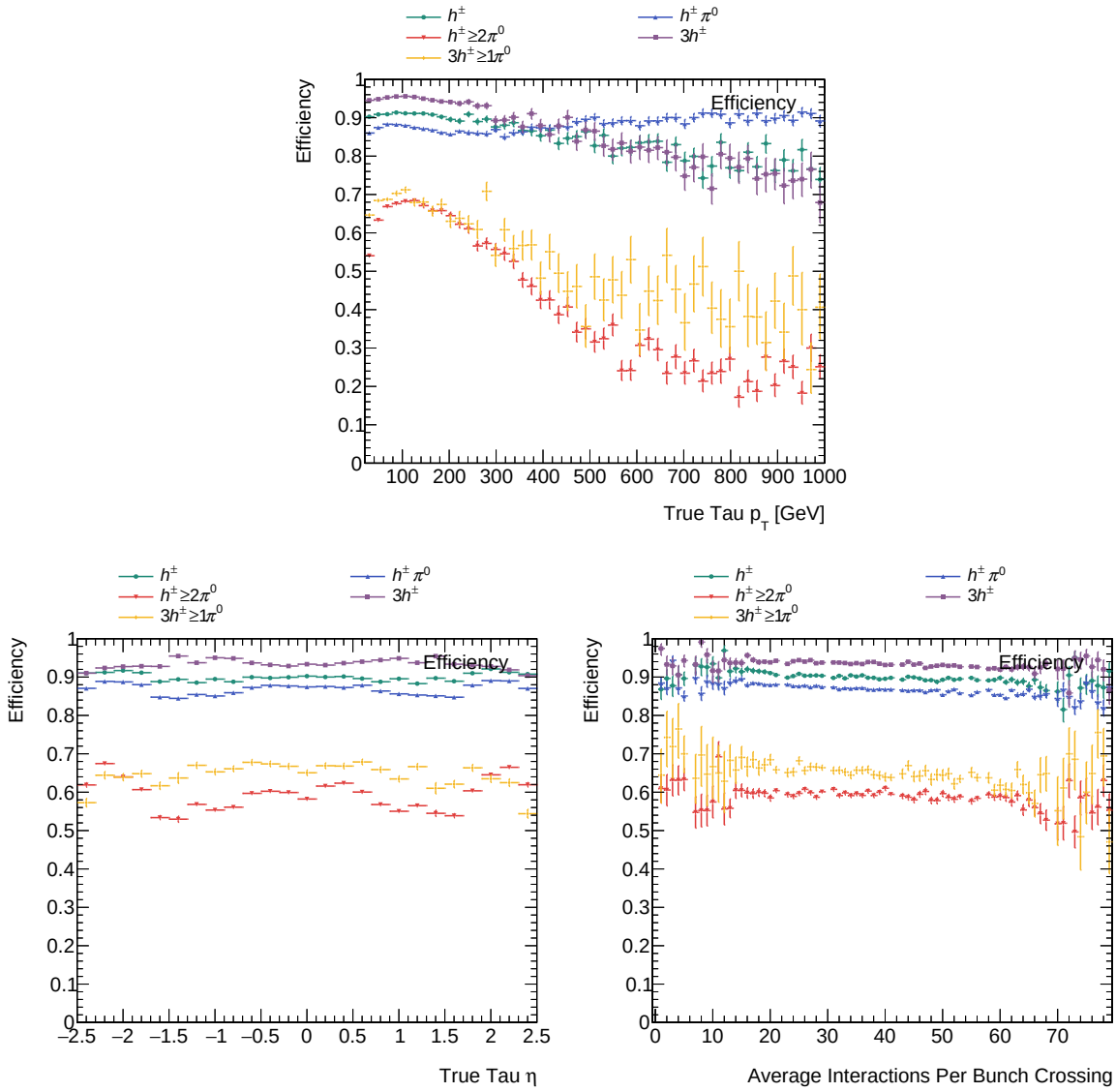


Figure A.2: Efficiency of identifying the correct decay modes as functions of the generator-level τ_{had} p_{T}, η and average number of bunch crossing for each generated decay modes. The error bar stands for the statistical uncertainty of simulated $\gamma^* \rightarrow \tau\tau$ samples.

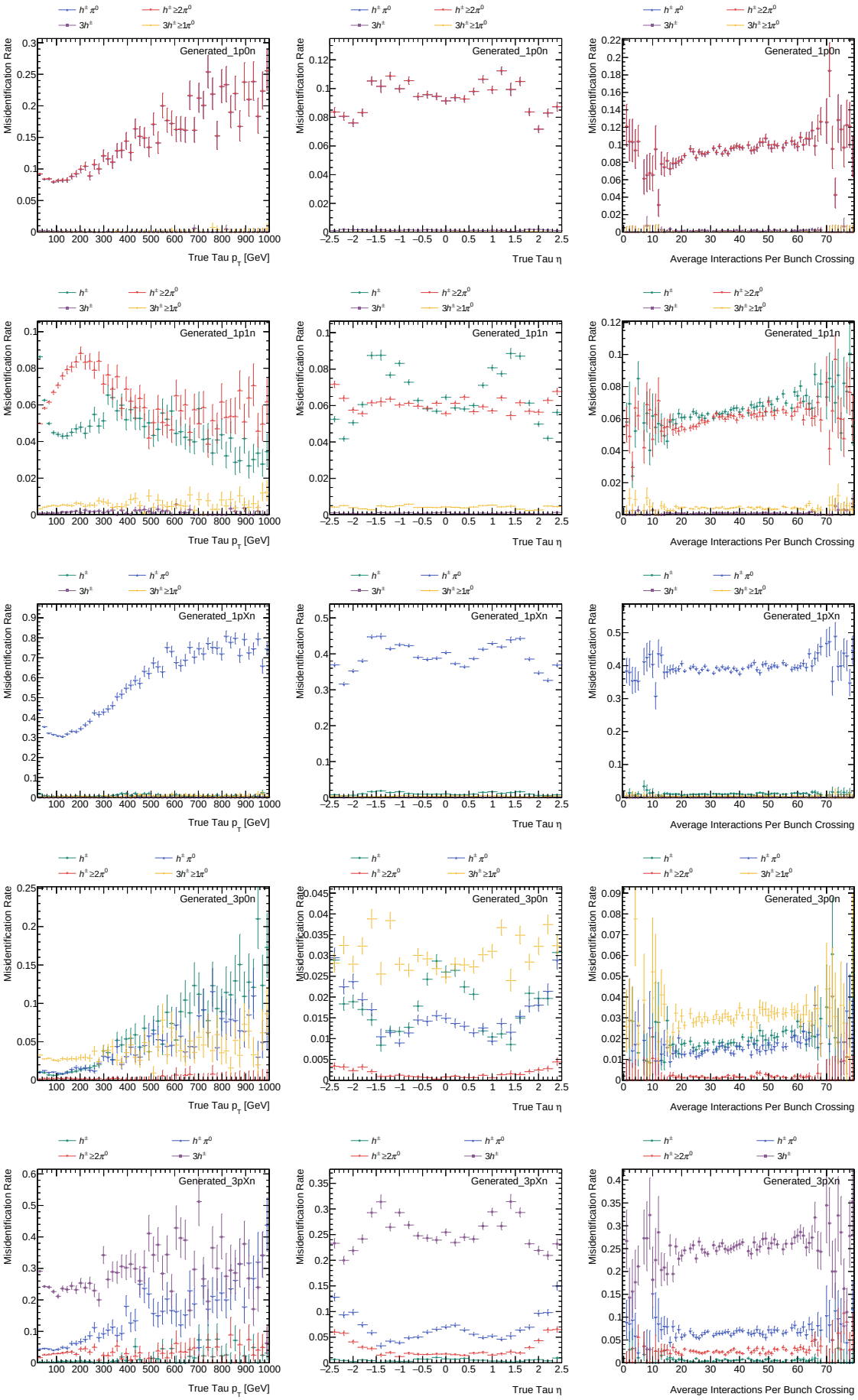


Figure A.3: Mis-identification rate of the decay modes as functions of the generator-level τ_{had} p_{T}, η and average number of bunch crossing for each generated decay modes. The error bar stands for the statistical uncertainty of simulated $\gamma^* \to \tau\tau$ samples.

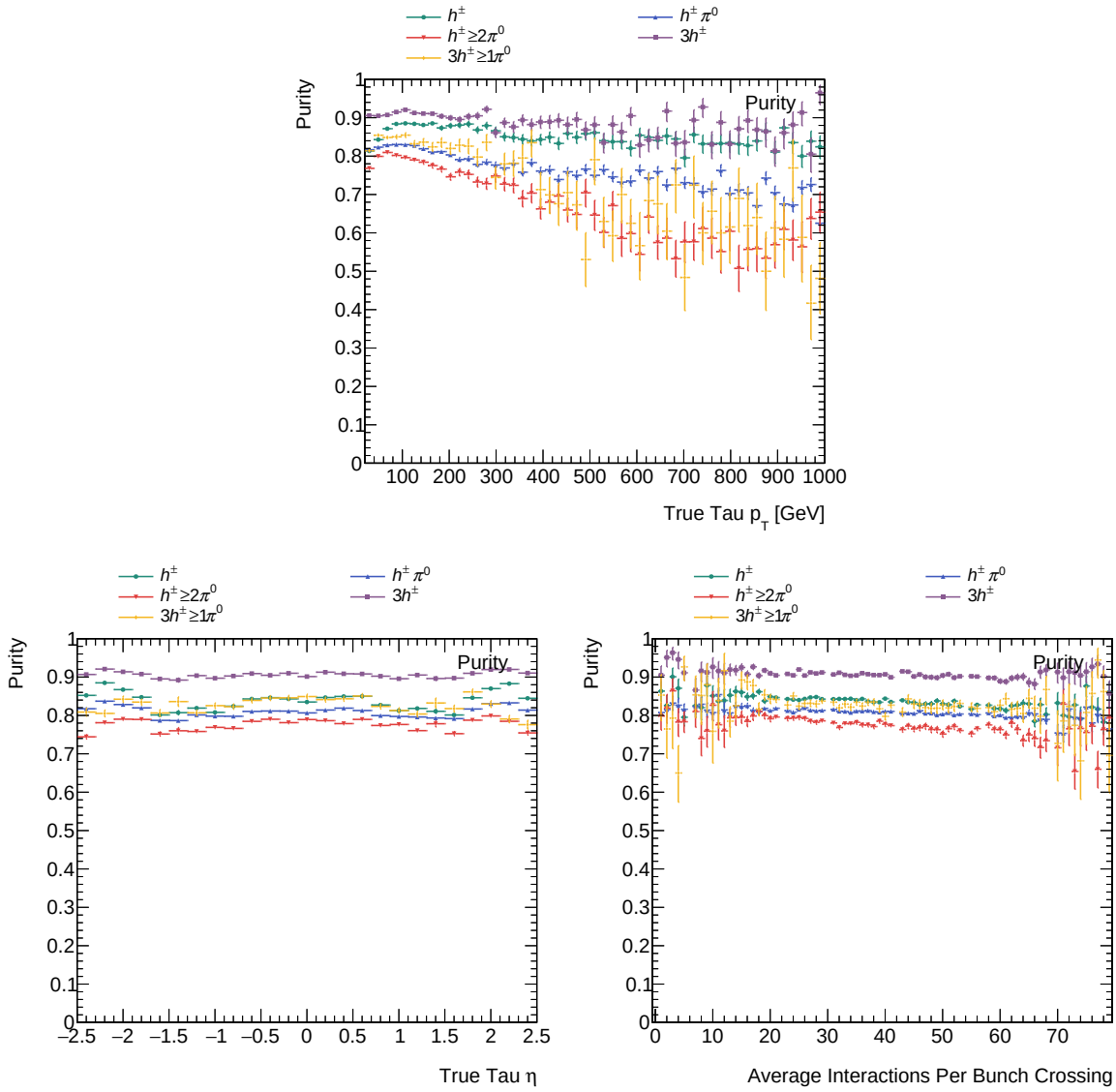


Figure A.4: Purity of the correctly identified decay modes as functions of the generator-level τ_{had} p_{T}, η and average number of bunch crossing for each generated decay modes. The error bar stands for the statistical uncertainty of simulated $\gamma^* \rightarrow \tau\tau$ samples.

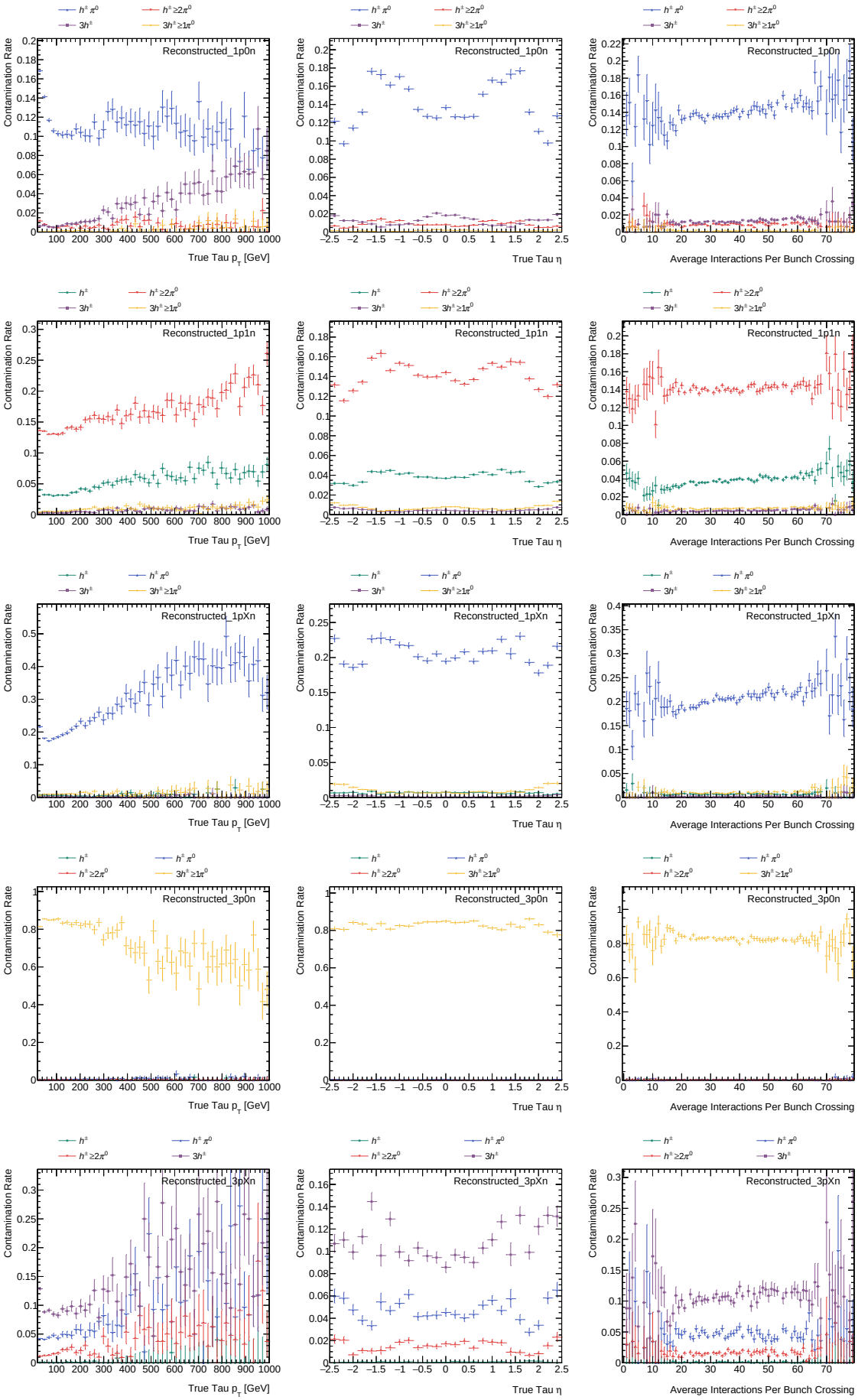
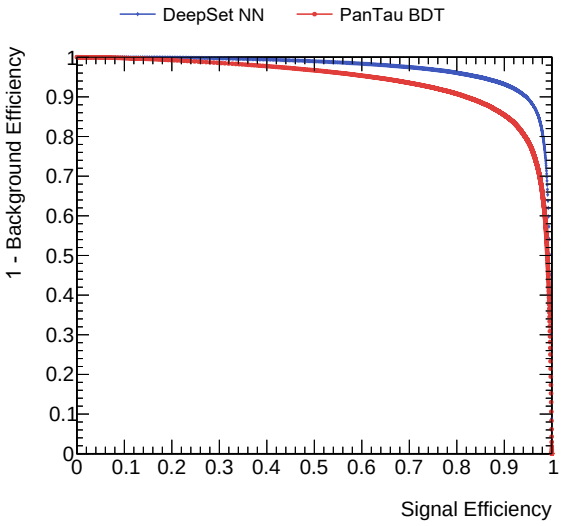
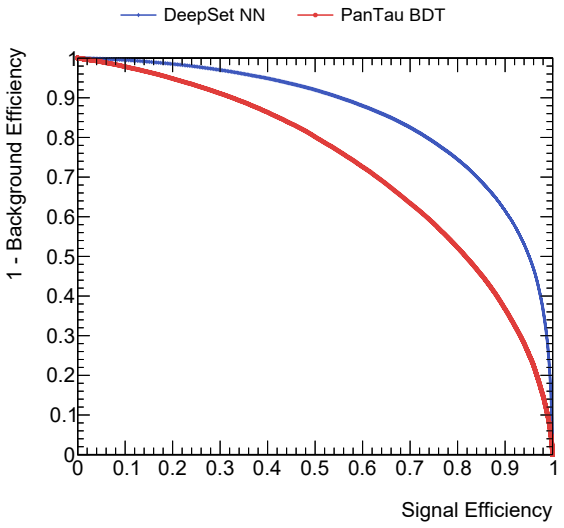


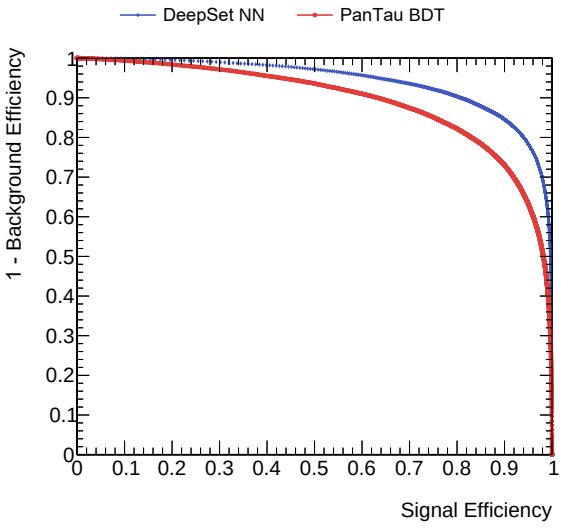
Figure A.5: Contamination rate of the decay modes as functions of the generator-level τ_{had} p_{T}, η and average number of bunch crossing for each generated decay modes. The error bar stands for the statistical uncertainty of simulated $\gamma^* \to \tau\tau$ samples.



(a) Signal:$h^{\pm}$, Background: $h^{\pm}\pi^{0}$



(b) Signal:$h^{\pm}\pi^{0}$, Background: $h^{\pm} \geq 2\pi^{0}$



(c) Signal:$3h^{\pm}$, Background: $3h^{\pm} \geq 1\pi^{0}$

Figure A.6: Relation between signal efficiency and 1− background efficiency of DeepSet (blue) and PanTau (red) decay mode classifiers in the discriminations of $h^{\pm}$ against $h^{\pm}\pi^{0}$, $h^{\pm}\pi^{0}$ against $h^{\pm} \geq 2\pi^{0}$ and $3h^{\pm}$ against $3h^{\pm} \geq 1\pi^{0}$ The error bar stands for the statistical uncertainty of simulated $\gamma^{*} \rightarrow \tau\tau$ samples.

Appendix B

Details on the search for Higgs boson pair production

B.1 Generalisation of MVA models

Figure B.1 confirms that there are no sign of overtraining by checking the compatibility of the MVA score distributions on training samples and testing samples in a binned Kolmogorov-Smirnov test (K-S test) [1]. The p-values of the tests are displayed in the legend. They are found to be close to 1 therefore the hypothesis of being compatible cannot be rejected.

B.2 MVA output distributions

The distribution of the non-resonant HH BDT and the resonant $X \rightarrow HH$ PNNs of twelve representative mass points after the likelihood fit under background-only hypothesis have been shown in Section 4.6. The distribution of the resonant $X \rightarrow HH$ PNNs for the rest of the mass points are shown in Figure B.2.

B.3 Validation of background estimation

This section summarised the validation of the background estimations described above. Experimental uncertainties are shown in the figures, but they will be explained in Section 4.4. The event yields and uncertainties shown in the figures are before the likelihood fits (pre-fit).

The Z+jets process is validated in the regions defined in Table B.1. The 2-b-tagged ZVR overlaps with the SR but the signals are excluded by the tight $m_{\tau\tau}^{\mathrm{MMC}}$ selection. Distributions of MVA input variables are shown in Figure B.3 and Figure B.4. Normalisation factors of 1.4 and 0.96 are applied to $Z + hf$ and $t\bar{t}$ process.

[1] The K-S test is implemented by ROOT [183]. 'X' option is used.

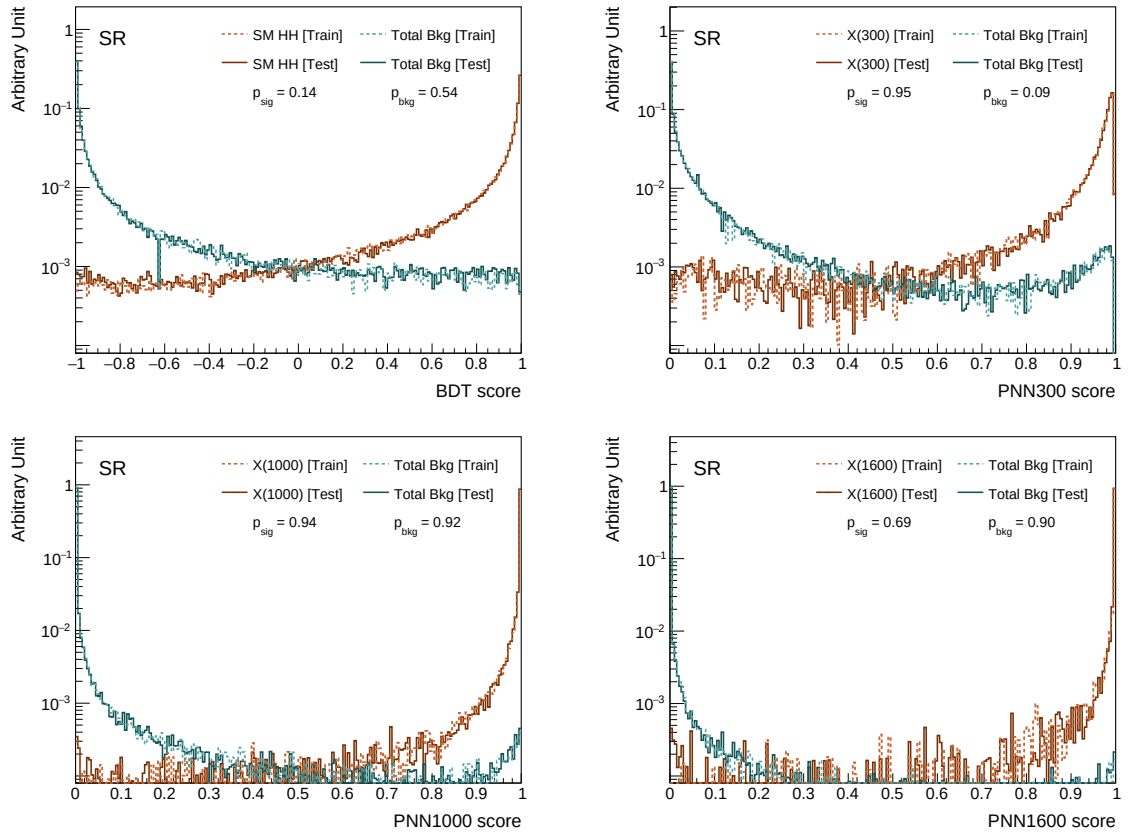


Figure B.1: Potential overtraining checked by testing the compatibility of the MVA score distributions on training samples and testing samples. The p-value of the hypothesis of being compatible is estimated by K-S test on signal and background classes are displayed on the legend of the figures as p_{sig} and p_{bkg}, respectively.

Table B.1: Definitions of Z+jets and $Z + hf$ validation regions.

Region name	Selections different from SR
1-b-tagged ZVR	$80 < m_{\tau\tau}^{\text{MMC}} < 100$ GeV, $\Delta R(\tau_{\text{had}}, \tau_{\text{had}}) < 1.5$, exactly one b-tagged jet
2-b-tagged ZVR	$80 < m_{\tau\tau}^{\text{MMC}} < 100$ GeV, $\Delta R(\tau_{\text{had}}, \tau_{\text{had}}) < 1.5$

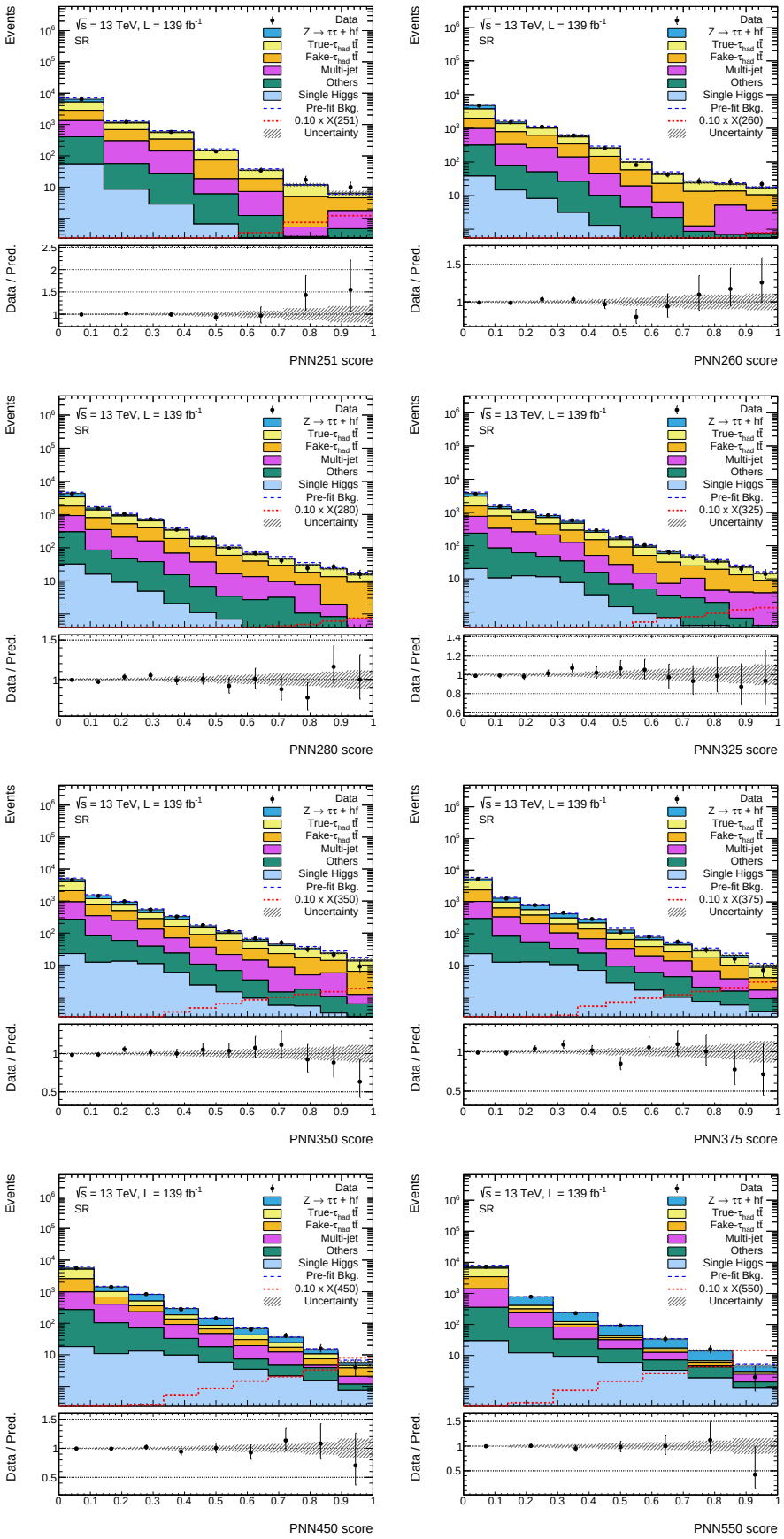


Figure B.2: Distributions of the PNN evaluated at m_X = 251-550 GeV for the resonant $X \to HH$ search after the corresponding background-only-hypothesis likelihood fits. The cross section of the resonant HH signal is assumed to be 1 pb, and the signals are scaled by a factor of 0.1. In the lower panel, the ratio of data and the total background prediction is displayed in black dots. The hatched area indicates the total systematic uncertainty of the backgrounds.

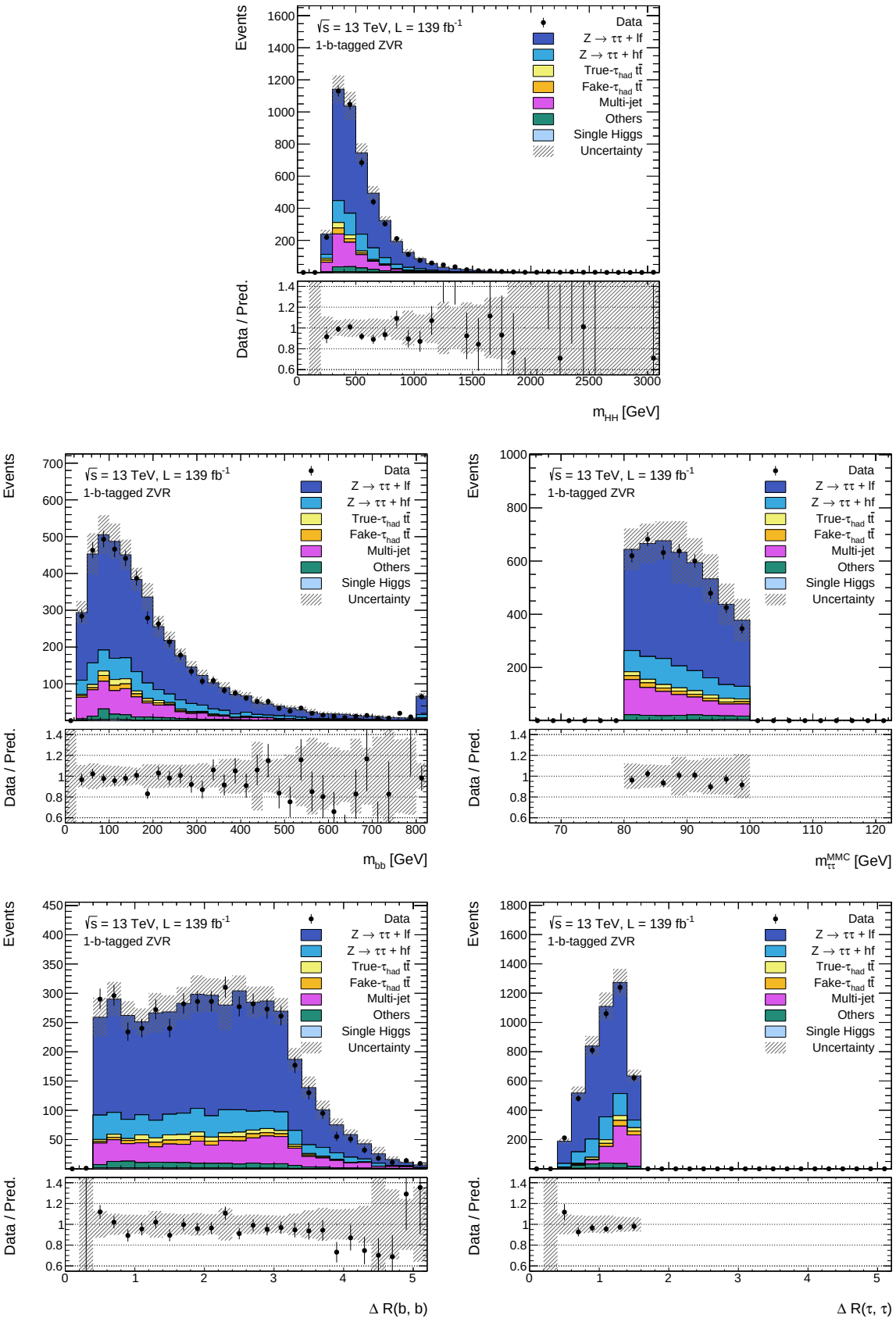


Figure B.3: Data and predicted background comparison in the 1-*b*-tagged ZVR. The uncertainty band includes the statistical uncertainty and the experimental uncertainty of the simulated sample. Overflow events are included in the last bin.

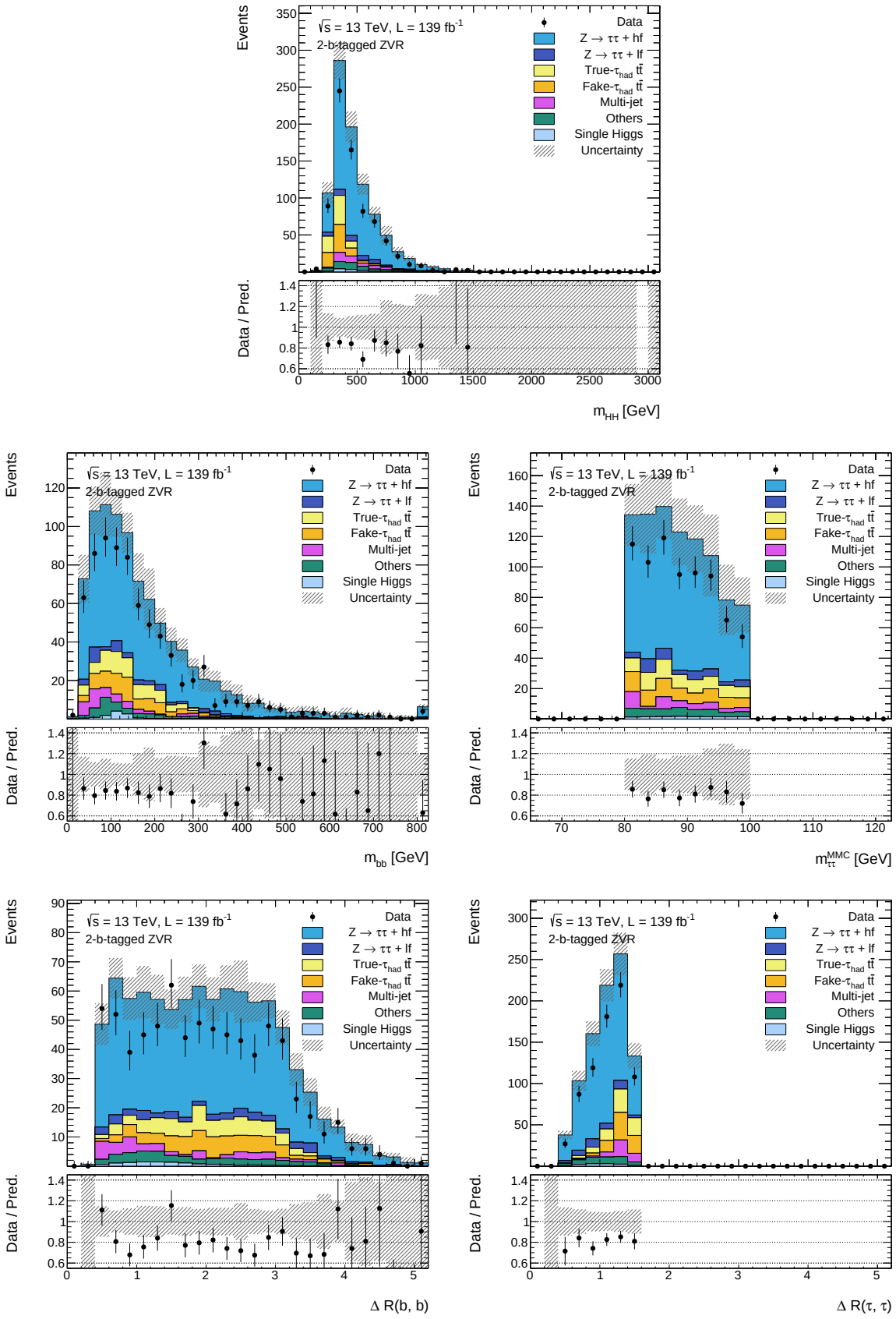


Figure B.4: Data and predicted background comparison in the 2-*b*-tagged ZVR. The uncertainty band includes the statistical uncertainty and the experimental uncertainty of the simulated sample. Overflow events are included in the last bin.

The $t\bar{t}$ reweighting is validated in $t\bar{t}$-VR1 and $t\bar{t}$-VR2. Their selections are listed in Table B.2. $t\bar{t}$-VR1 is defined using the same selection as $t\bar{t}$-noID-CR but invert the selection on m_{bb} (and veto the region near the mass of Higgs boson). $t\bar{t}$-VR2 requires *Loose* ID τ_{had} and $m_{\text{T}}^{W} > 140$ GeV, which is designed to exclusively check the performance of $t\bar{t}$ reweighting on true-τ_{had} $t\bar{t}$. Figure B.5 and Figure B.6 show the distribution of the p_{T} and jet multiplicity variables which are sensitive to $t\bar{t}$ reweighting.

Table B.2: Definitions of $t\bar{t}$ reweighting validation regions in $\tau_{\text{lep}}\tau_{\text{had}}$ final state.

Region name	Selections different from $t\bar{t}$-noID-CR
$t\bar{t}$-VR1	$m_{bb} \in (50, 80) \cup (140, 150)$ GeV
$t\bar{t}$-VR2	τ_{had} pass *Loose* ID, $m_{\text{T}}^{W} > 140$ GeV

The estimation of true- and fake-τ_{had} $t\bar{t}$ is inclusively checked in $t\bar{t}$-$\tau_{\text{had}}\tau_{\text{had}}$-VR in the $\tau_{\text{had}}\tau_{\text{had}}$ channel (see Table B.3). It also overlaps with SR but signals are excluded by the selections. Figure B.7 illustrates the distribution of MVA input variables. Normalisation factors of 1.4 and 0.96 are applied to $Z + hf$ and $t\bar{t}$ process.

Table B.3: Definitions of $t\bar{t}$ validation regions in $\tau_{\text{had}}\tau_{\text{had}}$ channel.

Region name	Selections different from SR
$t\bar{t}$-$\tau_{\text{had}}\tau_{\text{had}}$-VR	$m_{\tau\tau}^{\text{MMC}} > 110$ GeV, $m_{\tau\tau}^{\text{MMC}} + m_{bb} > 300$ GeV

The estimation of multi-jet is validated in several regions with different signs electric charges of τ_{had}'s and number of b-tagged jets, as summarised in Table B.4. Distributions of MVA input variables in these regions are shown in Figure B.8 and Figure B.9

Table B.4: Definitions of multi-jet validation regions.

Region name	Selections different from SR
1-b-tagged OS MJ VR	$m_{\tau\tau}^{\text{MMC}} > 110$ GeV, $E_{\text{T}}^{\text{miss}}$ significance < 3, exactly one b-tagged jet
2-b-tagged SS MJ VR	same-sign electric charges of τ_{had}'s

B.4 Global significance in resonant $X \to HH$ search

Figure B.10 is a demonstration of the correlation matrix described in Section 4.6.

Another bootstrap-based approach is developed to estimate the global significance for the local excess at 1 TeV resonance mass in the search for resonant $X \to HH$ search in the resolved analysis. The global significance derived in this approach needs to be treated with caution because the counting experiments ($\sum_i N^i_{\text{pseudo-data}}$) does not strictly follow Poisson distribution but a Skellam

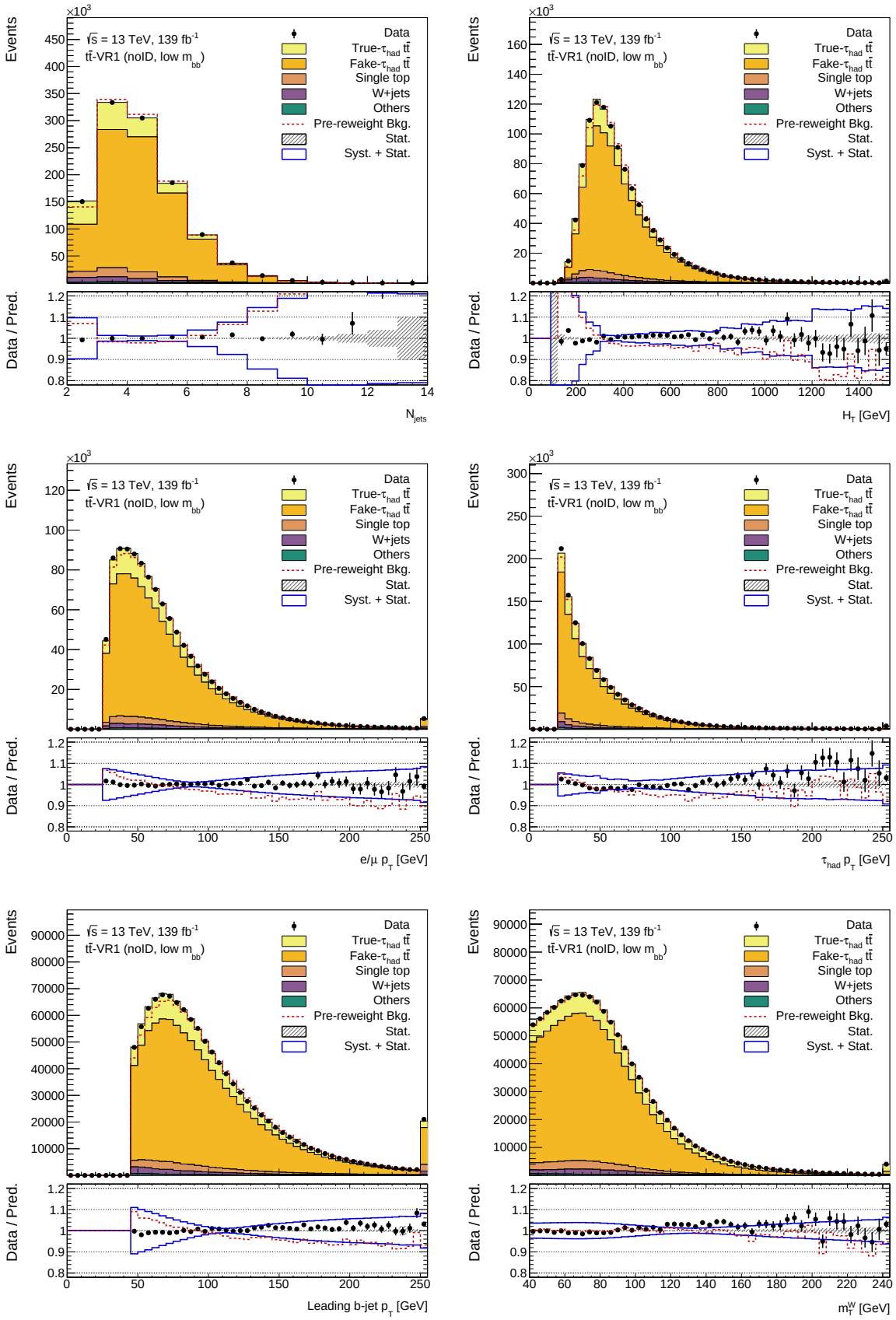


Figure B.5: Comparison of the kinematics of final state objects before and after $t\bar{t}$ reweighting. The shape of total background before reweighting (Pre-reweight) is displayed in dashed red line. The statistical uncertainty of simulated samples is shown in hatched boxes (Stat.), while the impact of systematic uncertainties of $t\bar{t}$ reweighting is indicated by the blue line (Syst. + Stat.). Overflow events are included in the last bin.

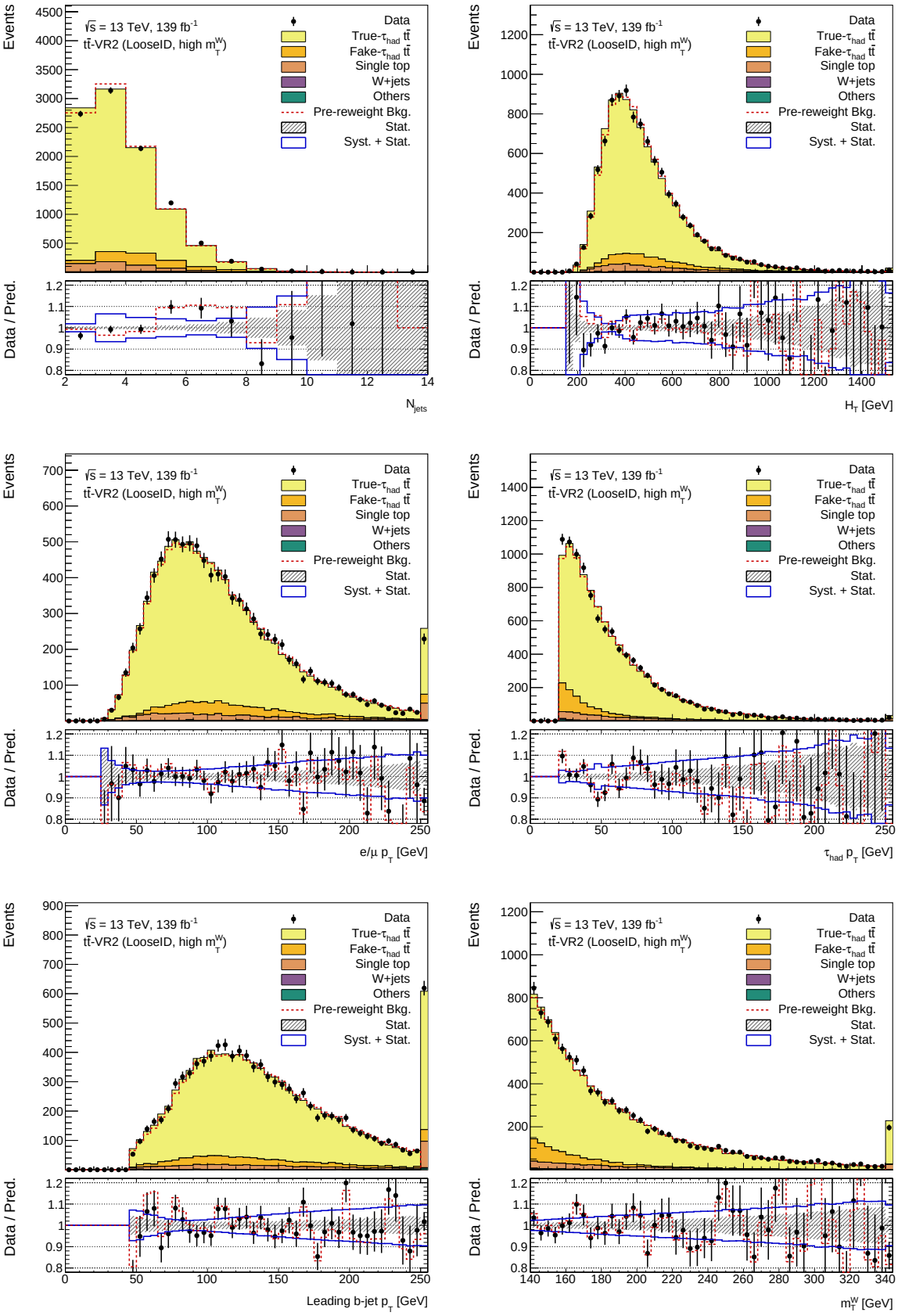


Figure B.6: Comparison of the kinematics of final state objects before and after $t\bar{t}$ reweighting. The shape of total background before reweighting (Pre-reweight) is displayed in dashed red line. The statistical uncertainty of simulated samples is shown in hatched boxes (Stat.), while the impact of systematic uncertainties of $t\bar{t}$ reweighting is indicated by the blue line (Syst. + Stat.). Overflow events are included in the last bin.

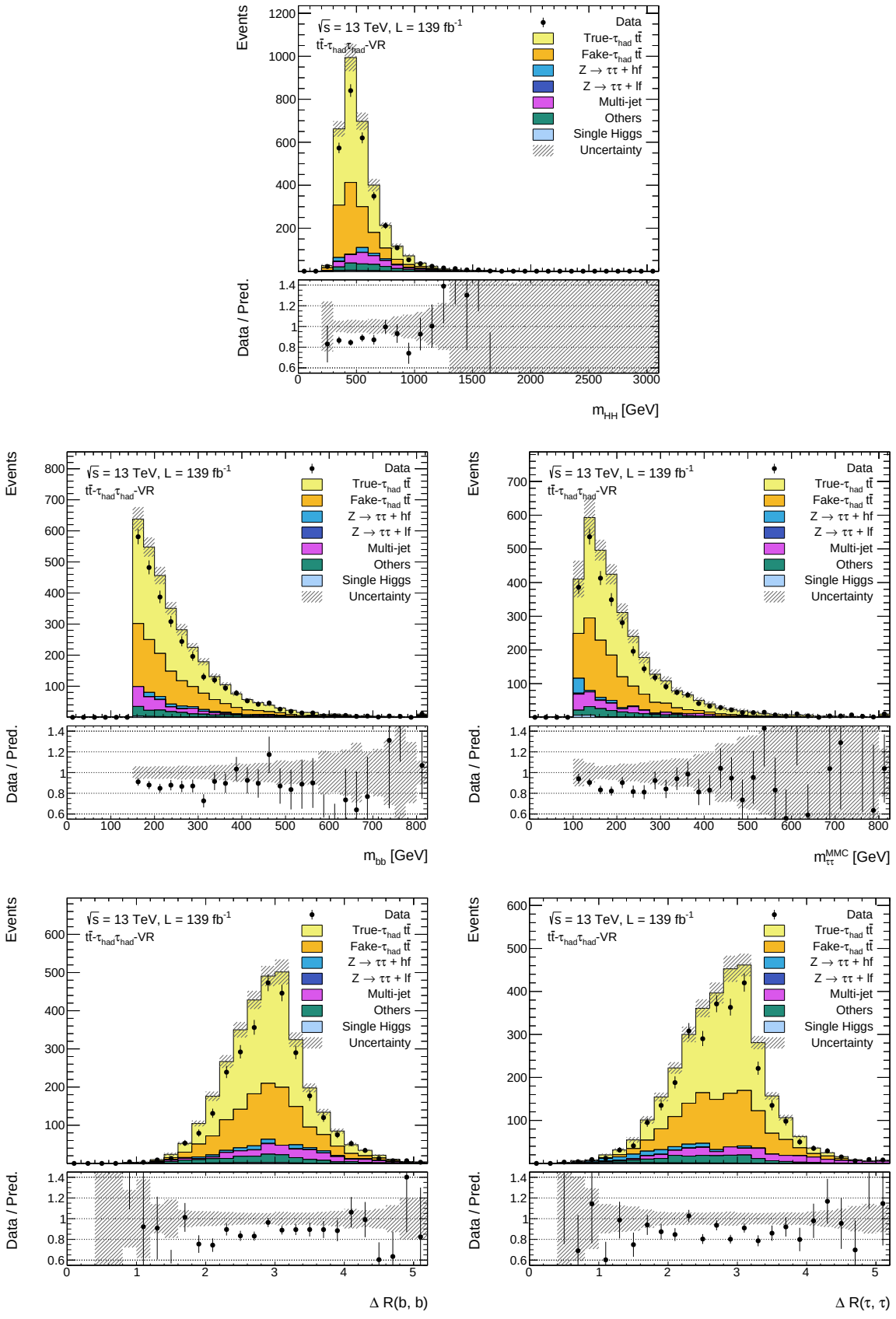


Figure B.7: Data and predicted background comparison in the $t\bar{t}$-$\tau_{\text{had}}\tau_{\text{had}}$-VR. The uncertainty band includes the statistical uncertainty and the experimental uncertainty of the simulated sample. Overflow events are included in the last bin.

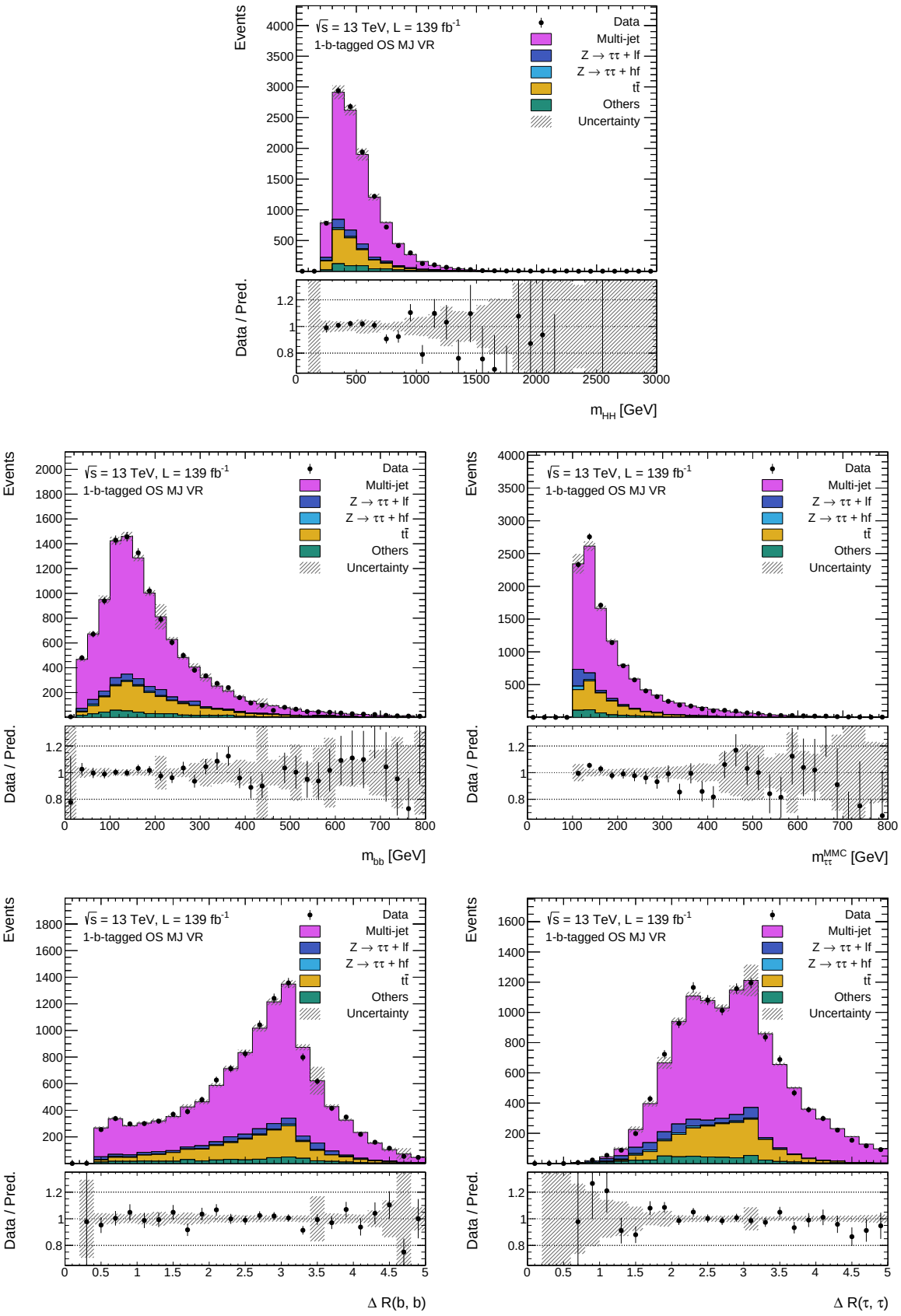


Figure B.8: Data and predicted background comparison in the 1-*b*-tagged OS MJ VR. The uncertainty band includes the statistical uncertainty and the experimental uncertainty of the simulated sample. Overflow events are included in the last bin.

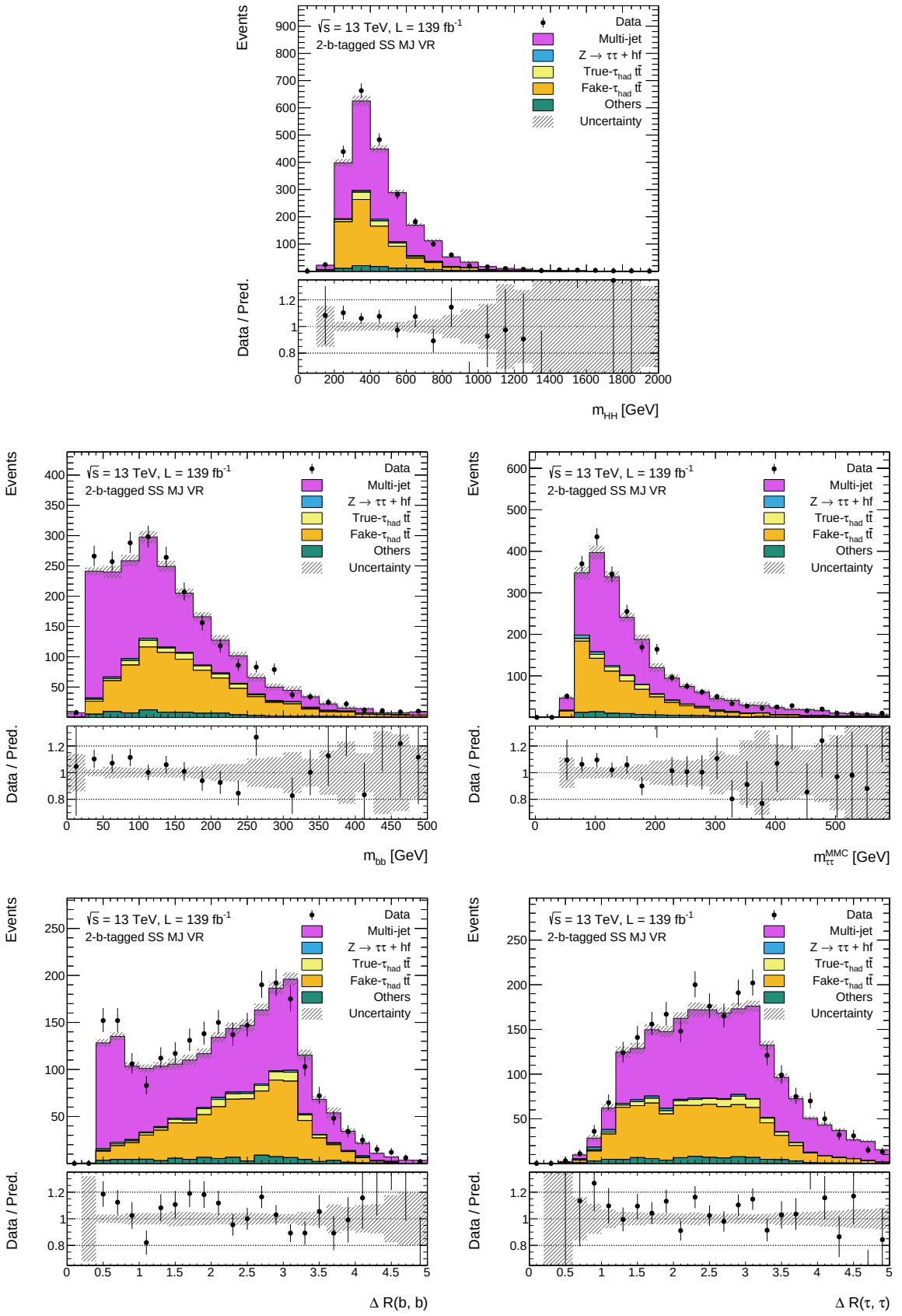


Figure B.9: Data and predicted background comparison in the 2-*b*-tagged SS MJ VR. The uncertainty band includes the statistical uncertainty and the experimental uncertainty of the simulated sample. Overflow events are included in the last bin.

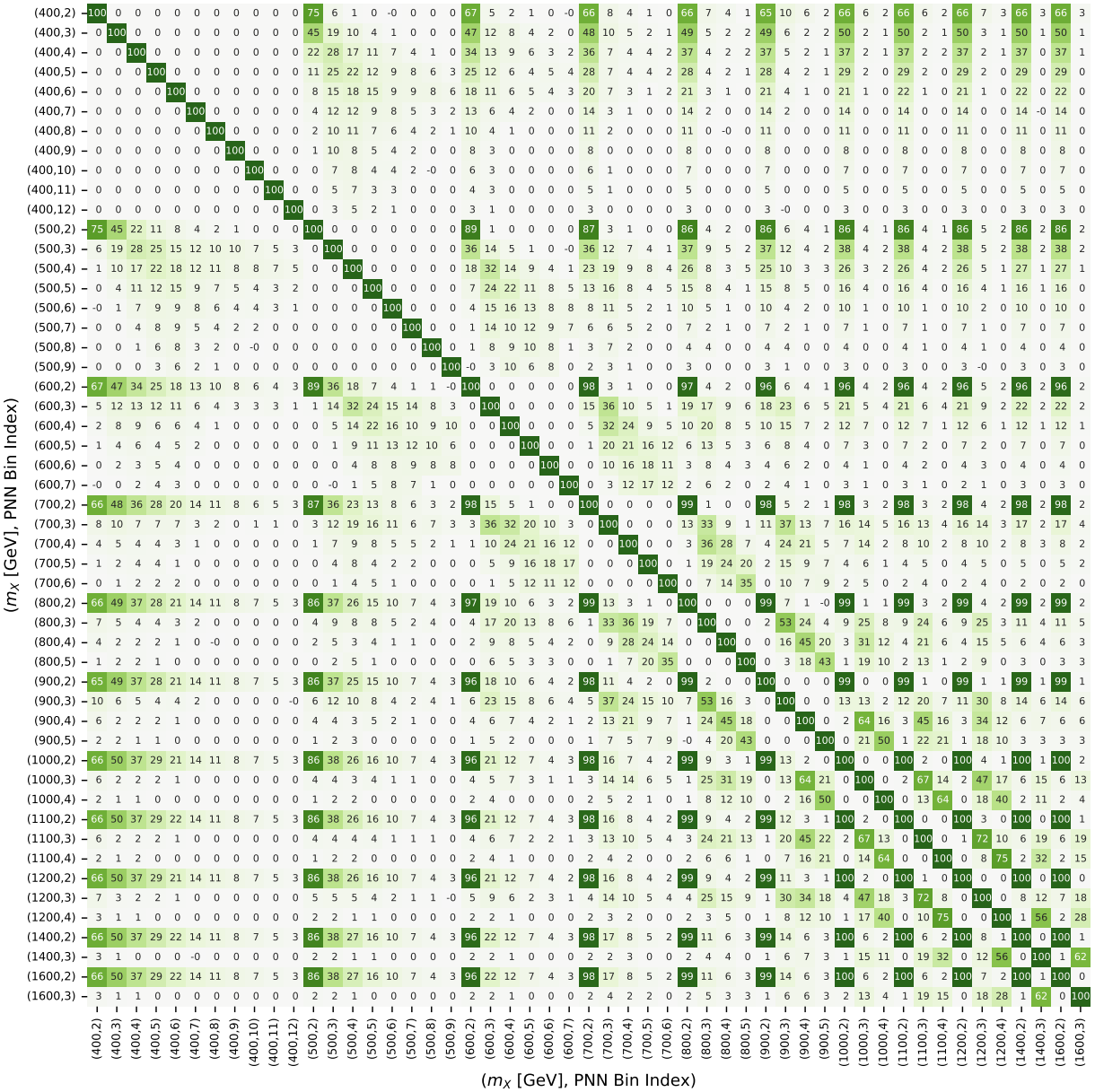


Figure B.10: Part of the correlation matrix R with indices labelled by mass hypothesis and PNN score bin index. The numbers shown are percentage values.

distribution. It is due to the existence of non-negligible negative weights in multi-jet events. However the two distributions only differ in variance, the result is still found to be a good estimation. Compared to the method described in Section 4.6, only the way that the pseudo-data are generated in the SR is different.

Here, the pseudo-data is sampled by a Poisson-bootstrap using the same post-fit expected background as the Gaussian copula approach, i.e., the yield of the i-th sampled pseudo-data ($N^i_{\text{pseudo-data}}$) is

$$N^i_{\text{pseudo-data}} = Poisson(w^i_{bkg}), \tag{B.1}$$

where w^i_{bkg} is the yield of the i-th weighted background event, $Poisson(\lambda)$ is a random variable generated based on parameter λ. For the multi-jet events estimated based on subtracting non-multi-jet simulated events from data, the weight can be negative. In that case, the absolute value is used and the sampled yield is subtracted. The correlation of the fluctuation of events in different search points is insured because the same sampled data is used in all mass hypothesis. The procedure is performed for both the SR and the $Z + hf$ CR [2]. The observables correspond to the auxiliary experiments are also resampled. Similar as the pseudo-data, the fluctuations of these observables are also correlated between different mass points. The same background and signal templates are used to calculate the local significances for each pseudo-experiment. Finally, the distribution of $Z^{\text{local}}_{\text{max}}$ is sampled, as shown in Figure B.11. The quoted uncertainty of $Z^{\text{glob.}}$ is due to the finite number of pseudo-experiments.

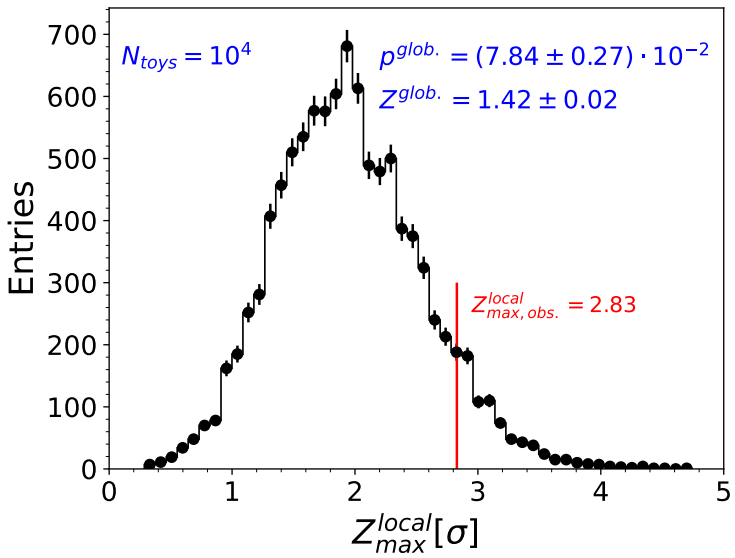


Figure B.11: Distribution of the maximum local significance of the 10^4 bootstrapped pseudo-experiments. The observed maximum local significance (at $m_X = 1000$ GeV mass point) is indicated in the red vertical line. The corresponding global p-value and significance are shown in blue text. The vertical error bar shows the statistical uncertainty of the pseudo-data.

[2] Since the same $Z + hf$ CR $m_{\ell\ell}$ distribution is used for all mass points. The sampling can be simplified.

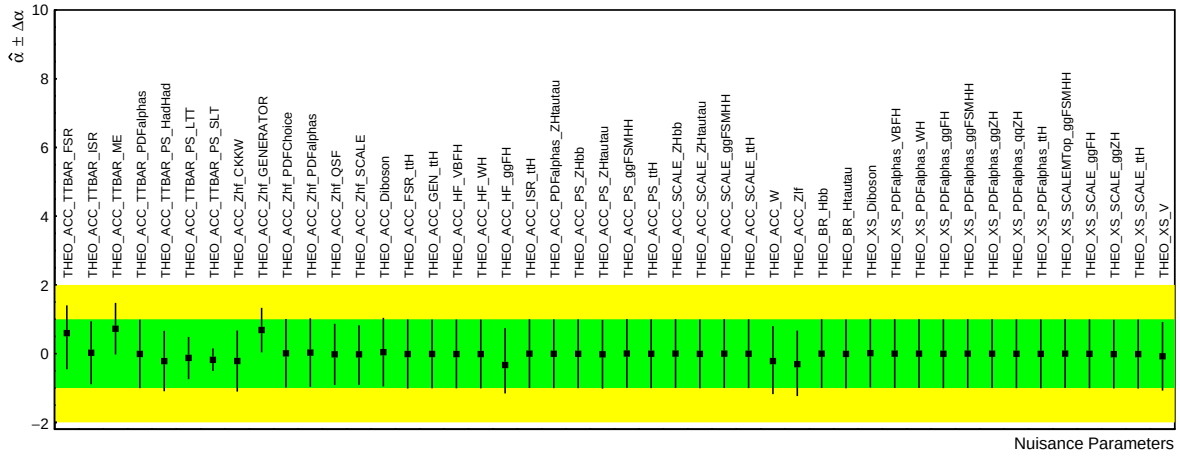
B.5 The combination of $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels

B.5.1 Nuisance parameters in the likelihood fit

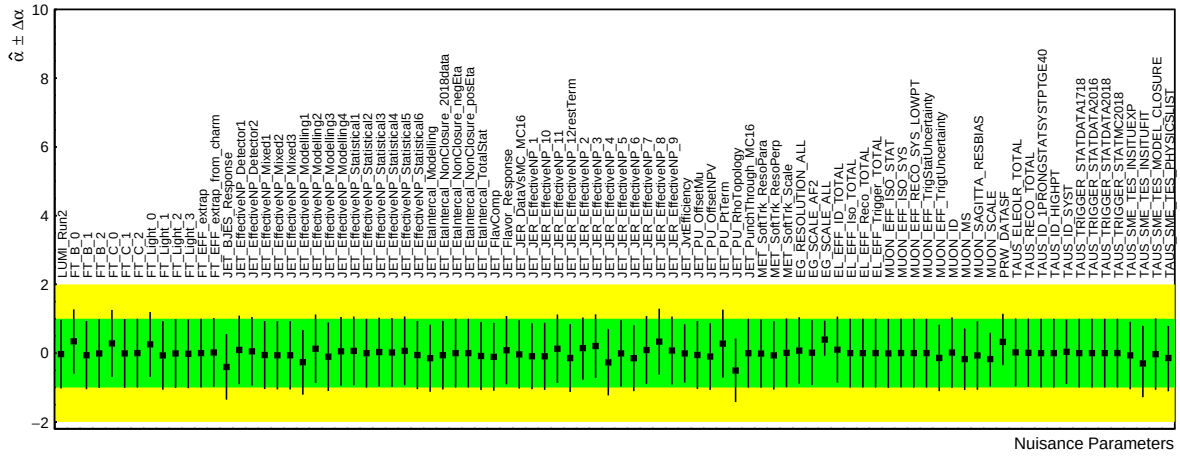
Several treatments are considered for numerical stability of the likelihood fit, which does not affect the results. The nuisance parameters that have negligible impact are ignored in the likelihood function. The $\pm 1\sigma$ histogram templates are smoothed by merging bins if large fluctuations present, before parameterized by the nuisance parameters. The estimators and 68% confidence intervals ($\hat{\alpha} \pm \Delta\alpha$) of the α nuisance parameters which are constrained by Gaussian terms in Equation (4.23), in the likelihood fit under background-only hypothesis combining non-resonant HH searches in the $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels, are shown in Figure B.12. The estimators and 68% confidence intervals of α parameters are found to be consistent with that of the corresponding auxiliary measurements (0 ± 1). The name of the nuisance parameters are indicated in Section 4.4, the ones in grey are nuisance parameters on the fake τ_{had} background estimation in $\tau_{\text{lep}}\tau_{\text{had}}$ [131]. The `PS` uncertainties in difference channels are uncorrelated, therefore three nuisance parameters are introduced for the uncertainty in $\tau_{\text{had}}\tau_{\text{had}}$, and `SLT` and `LTT` channels in $\tau_{\text{lep}}\tau_{\text{had}}$, as shown in their name. Figure B.13–B.16 shows the ranking of the nuisance parameters obtained in several resonant likelihood fits to data, to complement Figure 4.38–4.41.

B.5.2 Upper limits on resonant $X \to HH$ cross section

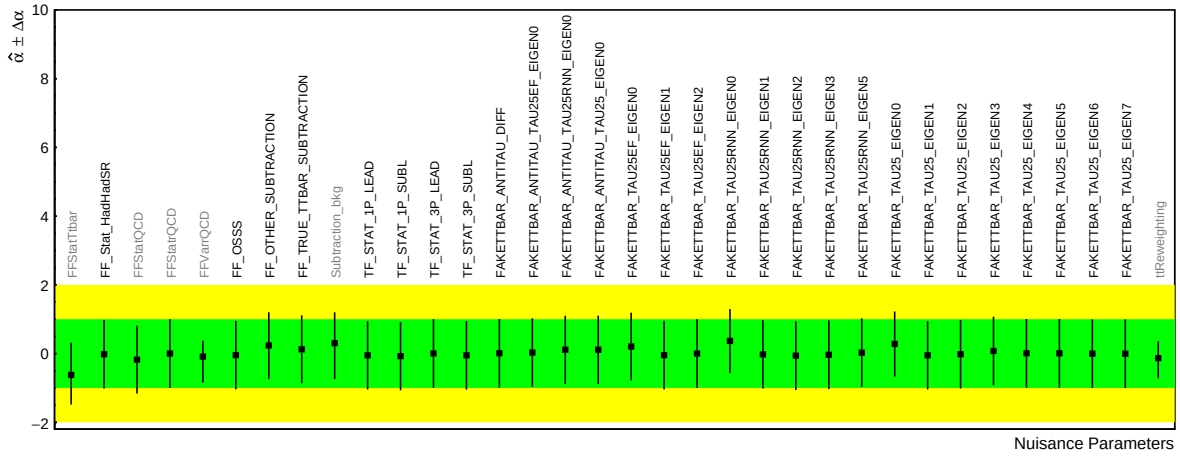
Figure B.17 summarises the upper limits on the resonant $X \to HH$ cross section in the resolved analysis (see Figure 4.43), the boosted analysis (see Figure 5.11), and the previous resolved analysis using the same final state [9].



(a) Theoretical



(b) Experimental



(c) Background estimation

Figure B.12: Estimators and 68% confidence intervals ($\hat{\alpha} \pm \Delta\alpha$) of the α nuisance parameters in the likelihood fit under background-only hypothesis combining non-resonant HH searches in the $\tau_{\text{had}}\tau_{\text{had}}$ and $\tau_{\text{lep}}\tau_{\text{had}}$ channels. These nuisance parameters are related to theoretical (a), experimental (b), and background-estimation (c) uncertainties.

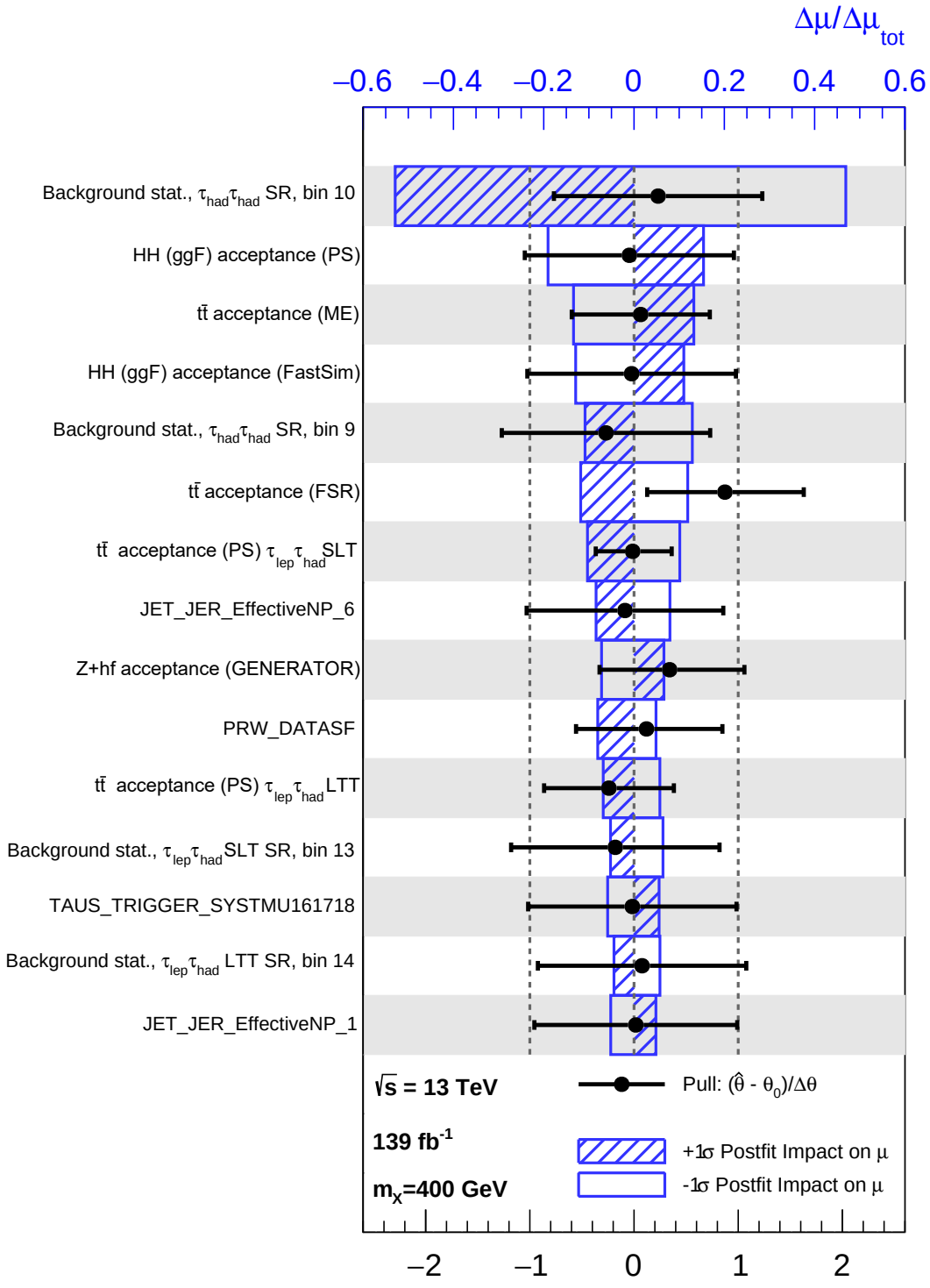


Figure B.13: Ranking of the nuisance parameters obtained in the $m_X = 400$ GeV resonant likelihood fits to data, ordered by the impact on the $\hat{\mu}$ divided by the confidence interval of $\hat{\mu}$ ($\Delta\mu/\Delta\mu_{\text{tot}}$). The $\pm 1\sigma$ impacts are expressed by hatched and and empty blue box, respectively. The black error bars shows the pulls of the nuisance parameters, which stands for the degree of deviation between the best-fit value and the value from the corresponding auxiliary measurements. The description of the nuisance parameters can be found in the text.

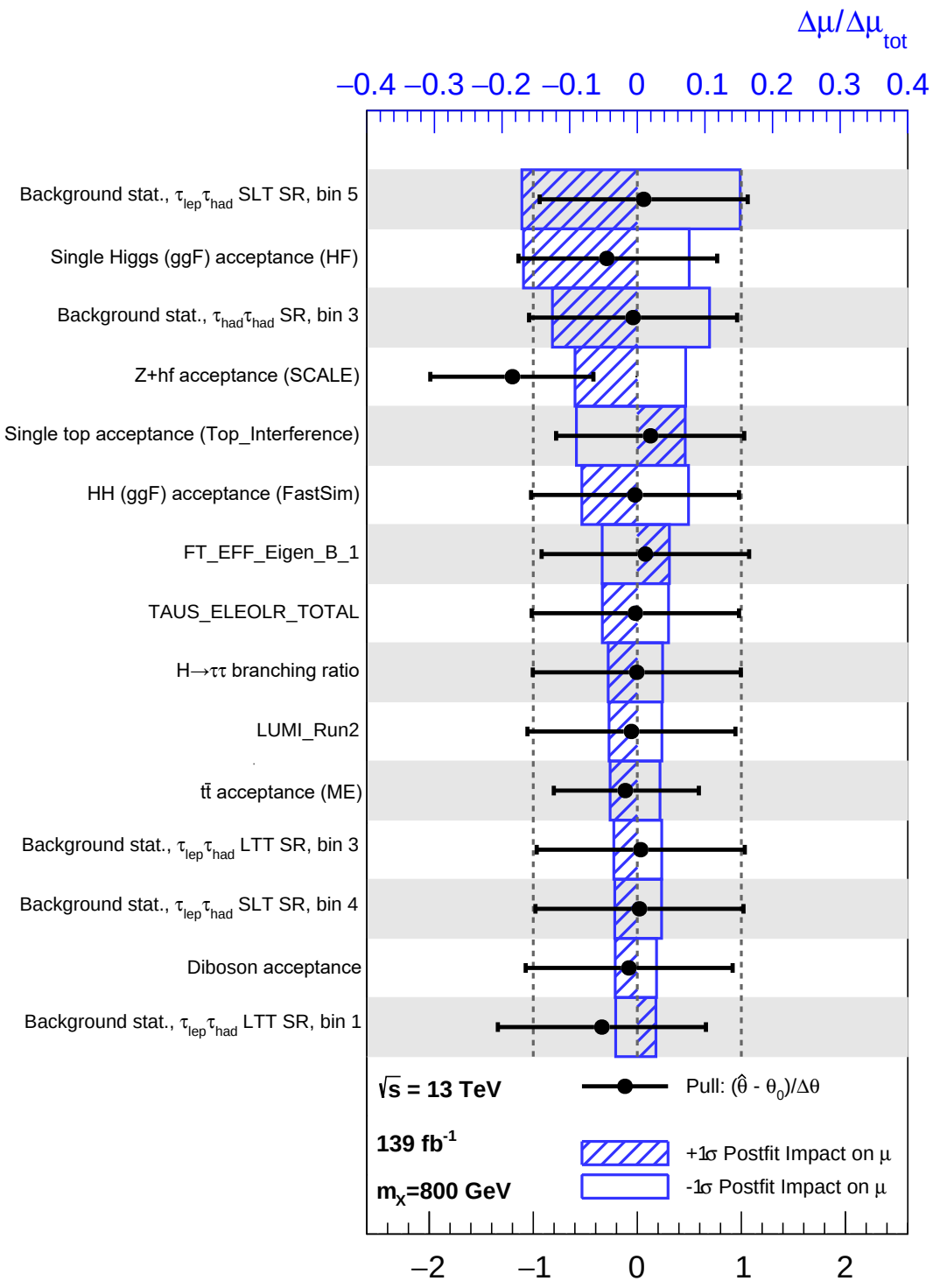


Figure B.14: Ranking of the nuisance parameters obtained in the $m_X = 800$ GeV resonant likelihood fits to data, ordered by the impact on the $\hat{\mu}$ divided by the confidence interval of $\hat{\mu}$ ($\Delta\mu/\Delta\mu_{\text{tot}}$). The $\pm 1\sigma$ impacts are expressed by hatched and and empty blue box, respectively. The black error bars shows the pulls of the nuisance parameters, which stands for the degree of deviation between the best-fit value and the value from the corresponding auxiliary measurements. The description of the nuisance parameters can be found in the text.

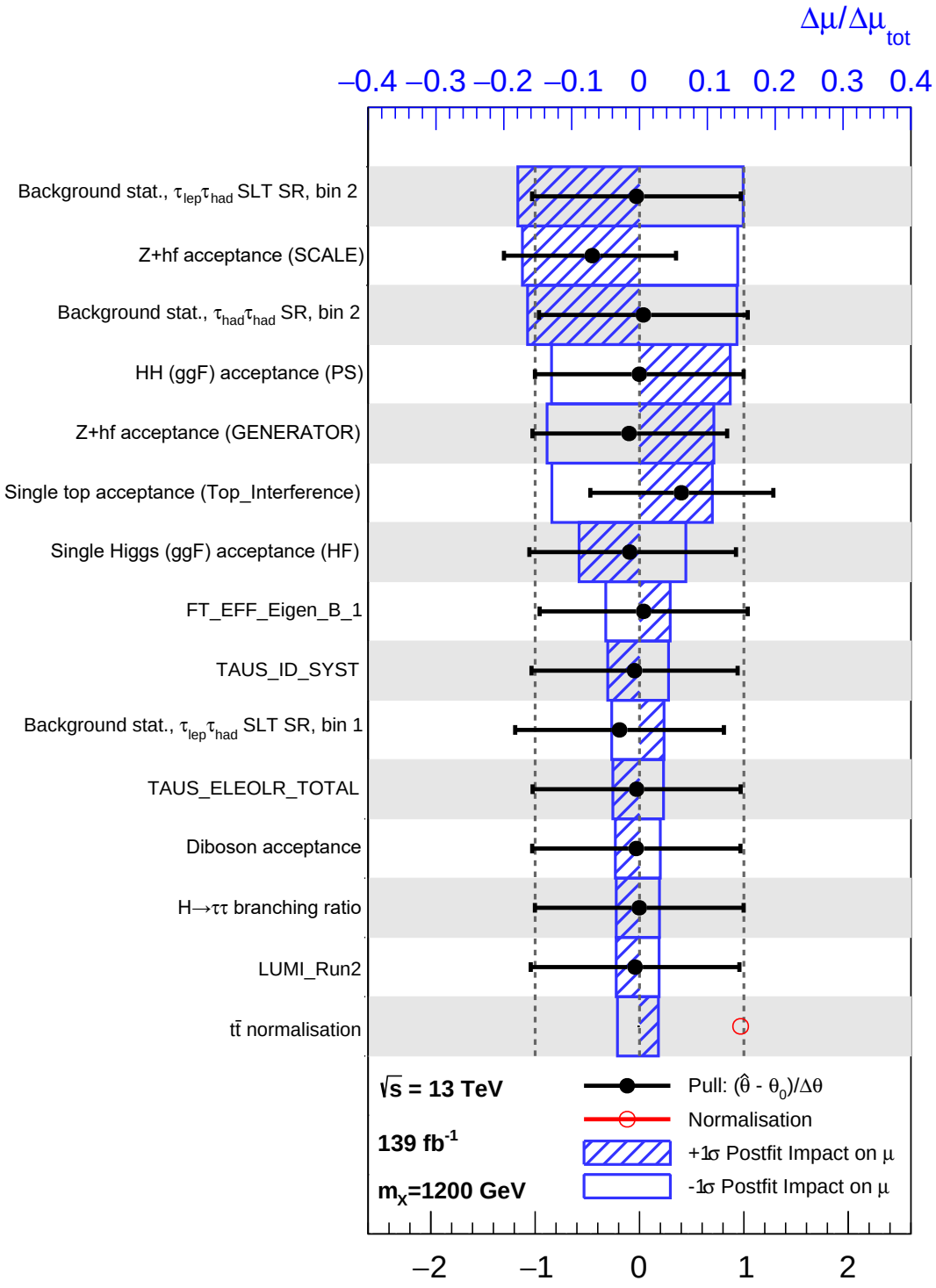


Figure B.15: Ranking of the nuisance parameters obtained in the $m_X = 1200$ GeV resonant likelihood fits to data, ordered by the impact on the $\hat{\mu}$ divided by the confidence interval of $\hat{\mu}$ ($\Delta\mu/\Delta\mu_{\text{tot}}$). The $\pm 1\sigma$ impacts are expressed by hatched and and empty blue box, respectively. The black error bars shows the pulls of the nuisance parameters, which stands for the degree of deviation between the best-fit value and the value from the corresponding auxiliary measurements. The description of the nuisance parameters can be found in the text.

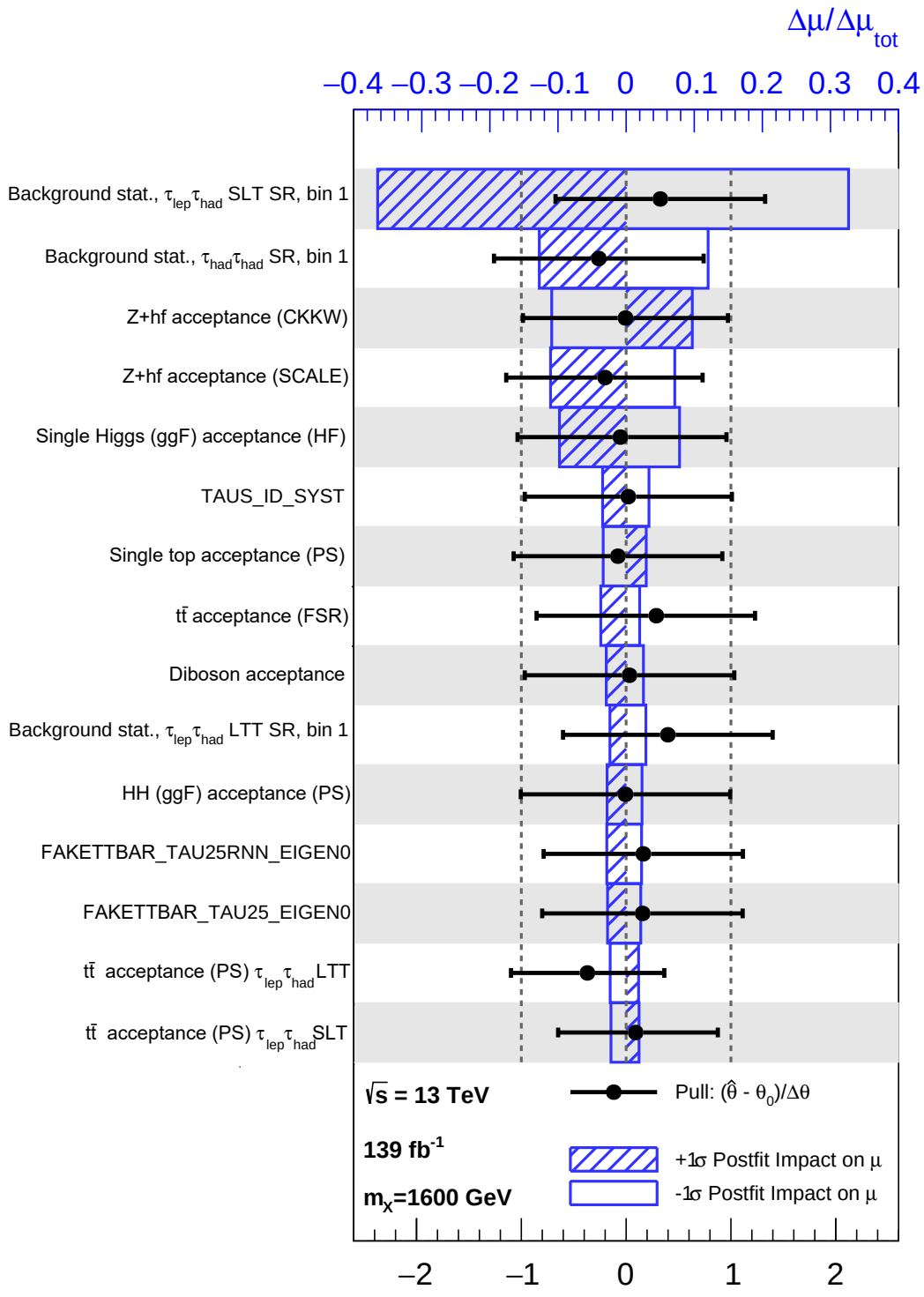


Figure B.16: Ranking of the nuisance parameters obtained in the $m_X = 1600$ GeV resonant likelihood fits to data, ordered by the impact on the $\hat{\mu}$ divided by the confidence interval of $\hat{\mu}$ ($\Delta\mu/\Delta\mu_{\text{tot}}$). The $\pm 1\sigma$ impacts are expressed by hatched and and empty blue box, respectively. The black error bars shows the pulls of the nuisance parameters, which stands for the degree of deviation between the best-fit value and the value from the corresponding auxiliary measurements. The description of the nuisance parameters can be found in the text.

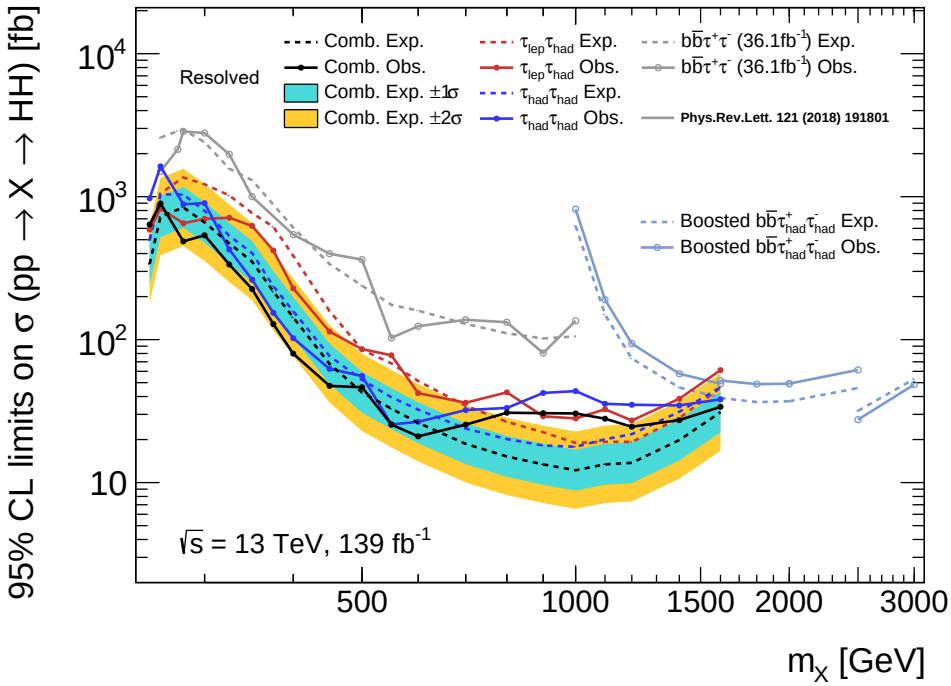


Figure B.17: Upper limits on the cross section of $gg \to X \to HH$ at 95% CL in the search for resonant Higgs boson pair production via a narrow-width scalar particle (X) as a function of m_X. Results from the resolved analysis (combination in black with $\pm 1/2\sigma$ error bands, $\tau_{\text{had}}\tau_{\text{had}}$ in blue and $\tau_{\text{lep}}\tau_{\text{had}}$ in red), the boosted analysis (greyblue), and the previous resolved analysis using the same final state [9] (grey) are overlaid. Observed and expected upper limits are indicated by the legends.

Publications and Conference Talks

Publications with significant contributions:

- ATLAS Collaboration, *Reconstruction and identification of boosted di-τ systems in a search for Higgs boson pairs using 13 TeV proton–proton collision data in ATLAS*, JHEP **11** (2020) 163.
- ATLAS Collaboration, *Search for resonant and non-resonant Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 TeV pp collision data from the ATLAS detector*, ATLAS-CONF-2021-030, 2021.

Publications with direct contributions:

- ATLAS Collaboration, *Combination of searches for non-resonant and resonant Higgs boson pair production in the $b\bar{b}\gamma\gamma$, $b\bar{b}\tau^+\tau^-$ and $b\bar{b}b\bar{b}$ decay channels using pp collisions at $\sqrt{s}$ = 13 TeV with the ATLAS detector*, ATLAS-CONF-2021-052, 2021.
- ATLAS Collaboration, *Validation of signal Monte Carlo event generation in searches for Higgs boson pairs with the ATLAS detector*, ATL-PHYS-PUB-2019-007, 2019.
- ATLAS Collaboration, *Summary of non-resonant and resonant Higgs boson pair searches from the ATLAS experiment*, ATL-PHYS-PUB-2021-031, 2021.
- ATLAS Collaboration, *Projected sensitivity of Higgs boson pair production in the $bb\tau\tau$ final state using proton-proton collisions at HL-LHC with the ATLAS detector*, ATL-PHYS-PUB-2021-044, 2021.

Conference talks:

- **International conference:** *Searches for new resonances decaying to HH at ATLAS*, Higgs 2020 Conference, Stonybrook/BNL, Oct 2020.
- **National conference (poster):** *Search for Higgs boson pair production using τ-related final states in ATLAS*, The 5th China LHC Physics Conference, 大连, Oct 2019.
- **National conference:** *Search for resonant Di-Higgs production using boosted $b\bar{b}\tau^+\tau^-$ final state at ATLAS*, The 6th China LHC Physics Conference, 北京, Oct 2020.
- **National conference:** *Search for Higgs boson pair production in the $b\bar{b}\tau^+\tau^-$ decay channel using 13 TeV pp collision data from the ATLAS detector*, 中国物理学会高能物理分会第十三届全国粒子物理学术会议, 青岛, Aug 2021.

Acknowledgement

致谢

论文的完成首先要感谢我的导师陈申见老师。从本科的粒子物理课开始，就被陈老师丰富的学识和儒雅的气质吸引。陈老师对粒子物理实验研究方向独特而深刻的见解带给我很多启发。生活中，陈老师风趣和蔼，常关心我和同学们的生活情况和未来发展。在科研之路初始，陈老师耐心引领我入门，让我逐渐对粒子物理实验的各个方面形成了初步的认识。后来，陈老师鼓励我开拓视野，不局限于自己的工作内容，并培养自己提出问题解决问题的能力。在最后的论文写作阶段，陈老师给予了很多宝贵建议，并细致地指导我提高论文的质量。

另外还要感谢我的导师张雷老师。张老师认真积极的态度和对粒子物理的热忱既是我的榜样，也是我努力工作的动力。和张老师的日常讨论带给我对许多问题更多的思考和灵感。初入 ATLAS 实验时，张老师提供了非常耐心而具体的指导，在技术上和实验方法上传授了很多宝贵的经验。张老师把我介绍给了 tau 工作组和 di-Higgs 工作组，也因此才有此论文的成果。由衷地感激张老师为我能够在合作组顺利开展研究而付出的沟通协调工作。最后对论文的结构和内容，张老师也给予了很多具体的建议。

在 ATLAS 实验中，我遇到了很多来自各地的老师和同事，他们对我的科研工作提供了诸多具体的建议，和他们的讨论总能让人打开思路。感谢所有 tau 工作组召集人，他们为我提供了软件以及 tau 衰变分类等项目的工作机会，并且持续提供宝贵建议。感谢所有 di-Higgs，combination 和 bbtautau 工作组召集人，他们为我提供了 bbtautau 和 multi-lepton 等工作机会，并在分析研究的各个阶段都给予了很多支持和建议。还要感谢 Tina，Theodore，David，Alessandra，Chris 等来自其他机构的同事，在合作中他们给了我许多帮助。特别是 David，非常感谢由他帮忙，由 TU Dresden 慷慨提供的计算资源。

在南京大学高能物理组，我有幸结识了很多优秀的同学，在同他们的讨论中我一次次深受启发。晓忠和 Antonio 两位博士后总能轻松化解我的疑问。瀚非和汇润同学让我了解了很多研究课题以外的方法和思想。还感谢他们，以及亮亮同学，和我一起讨论论文内容和准备答辩，在此期间一起度过了很多快乐的时光。感谢其他同窗，开始时 523 办公室的友华，王蔚，Fabio，后来 526 办公室的立民，成伟，国君，昱昊，子航，振武，谢陈，梓峰，桓国，轶旻，以周，以及同组的家梁，王新，等等。难忘每一次知识和故事的分享。

最后感谢我的家人，我亲爱的梁冉和我的朋友们，他们的理解的支持让我可以安心的投入科研并完成论文。我想把论文献给离世不久的祖父，他永远是我心中的英雄。